AD_____

Award Number: W81XWH-05-1-0314

TITLE: Early Life Processes, Endocrine Mediators, and Number of Susceptible Cells in Relation to Breast Cancer Risk

PRINCIPAL INVESTIGATOR: Dimitrios Trichopoulos

CONTRACTING ORGANIZATION: Harvard College
Boston, MA 02115

REPORT DATE: April 2011

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*

4. TITLE AND SUBTITLE	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) E-Mail:	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU		**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Page

INTRODUCTION

The aim of the project was to investigate the role of early life processes, endocrine mediators and number of susceptible cells on adult life breast cancer risk. Based on the hypothesis that breast cancer risk is a function of number of mammary gland cells at risk of transformation and that this number is largely modulated by perinatal events and conditions, five component projects were initiated. The first three focused on perinatal characteristics, including immediate postnatal growth, in relation to mammary gland mass and breast cancer risk, whereas the last two explored the relation of pregnancy hormones with breast cancer risk and with cellular populations that are likely to correlate with mammary stem cell potential. The five projects were interlinked and they addressed the hypothesis that growth and mammotropic hormones in perinatal life affect the number of susceptible mammary gland cells. This number is likely to be reflected in birth size and rate of postnatal growth that, in turn, represent intermediate steps and correlates of mammary gland mass and breast cancer risk in adult life. The progress on each component project (CP) will be reported separately to facilitate the reader.

BODY

CP1 "Association of growth during the first postnatal week with breast cancer risk in adult life"

CP1 PI: Prof. Anders Ekbom, Unit of Clinical Epidemiology, Dept. of Medicine, Karolinska Institutet/Karolinska University Hospital, SE-171 76 Stockholm, Sweden.

Timetable of research accomplishments of CP1 as outlined in the Statement of Work.

Task 1 *To investigate the association of growth during the first postnatal week with breast cancer risk in adult life:*

a. Retrieval of available birth records from 1,068 women with incident breast cancer and 2,727 control women. (Months 1-24)
b. Extraction of data on growth of newborns during the first postnatal week, as well as information on covariates to be used in the analysis. (Months 25-30)
c. Linkage of data on postanatal growth and perinatal covariates to cancer and mortality registries. (Months 31-36)
d. Data analyses. (Months 37-48)
e. Manuscript preparation and submission. (Months 49-60)

CP1 progress report

The retrieval of birth records (task 1a), as well as the extraction of data on growth of newborns (task 1b) and linkage with cancer and mortality registries (task 1c) have been completed. In accordance with the time plan, tasks 1d and e have now been completed.

Specifically:

- The retrieval of birth records from 1,068 women with incident breast cancer and 2,727 control women has been completed.
- The extraction of data on growth of newborns during the first postnatal week has been completed. Birth records with postnatal growth information were available for 405 breast cancer patients and 1081 age- and hospital-matched controls.
- Linkage with cancer and mortality registries has been completed.
- Data analyses have been completed. A manuscript has been published.

CP1 key research accomplishments

Summary
We found that postnatal weight loss (an indicator of water loss, likely to reflect water retention associated with pregnancy hormones), as well as neonatal weight gain rate after reaching the nadir (known to reflect growth hormone levels), are significantly positively associated with premenopausal breast cancer risk (Lagiou et al, 2008 – see Appendix 2)

Specifically:
Birth size has been positively associated with breast cancer risk in the offspring (PLoS Med 2008; 5: e193), but no study had evaluated the association of neonatal growth with this risk, even though neonatal growth could be of particular importance, as it is strongly associated with neonatal IGF-1 levels (Paediatr Perinat Epidemiol 2003;17:281-6). IGF-1 levels, which could track through life, have been associated with breast cancer risk, particularly premenopausal breast cancer risk (Lancet 2004;363:1346–1353 ; Lancet Oncol. 2010;11:530-42.).

We conducted a case-control study nested within a population-based cohort of women, born in Sweden from 1901 through 1961. Newborn charts were available for 405 of the 1068 eligible breast cancer patients and for 1081 of the 2727 eligible control women, all born from 1901 and on. The general growth pattern of newborns is a decline in weight during the first week of the post-natal period, followed by an increase in weight (Arch Dis Child Fetal Neonatal Ed 2003;88:F472-6). To examine whether these two different phases of postnatal pattern of growth were associated with subsequent risk of breast cancer, we determined maximum postnatal weight loss [defined as (birth weight) – (lowest weight recorded during the hospital stay)] and the rate of growth since the nadir [defined as (weight at discharge – weight at nadir)/(day at discharge - day at nadir)].

We created five mutually exclusive categories: (a) neonates who remained at the maternity wards for more than 21 days without regaining their birth weight - these neonates were analyzed separately, because their weight loss and gain did not conform to a usual pattern; (b) neonates with a maximum postnatal weight loss of <200g and rate of growth after nadir <25g/day; (c) neonates with a maximum postnatal weight loss of >=200g and rate of growth after nadir <25g/day; (d) neonates with a maximum postnatal weight loss of <200g and rate of growth after nadir >=25g/day; and (e) neonates with a maximum postnatal weight loss of >=200g and rate of growth after nadir >=25g/day. The weight loss cutoff of 200 g was a round figure derived from the 6.6% that has been reported to be the median percent of birth weight loss (Arch Dis Child Fetal Neonatal Ed 2003;88:F472-6), so that with birth weight

around 3000 g, we have 3000 g * 0.066 ≃ 200g. The cut-off for the daily rate of growth after nadir was rounded at 25g/day, taking into account that the median time for birth weight recovery among children has been reported to be 8.3 days (Arch Dis Child Fetal Neonatal Ed 2003;88:F472-6), so that 200g divided by 8.3 days equals approximately 25g/day.

Statistical analyses were undertaken to examine postnatal growth patterns in relation to breast cancer risk by modelling the data through conditional logistic regression. Covariates adjusted in the analysis included maternal age (in years, as a continuous variable), maternal socioeconomic status (low, medium, and high, as an ordinal variable), maternal parity (1, 2, and >=3, as categorical indicator variables), pregnancy toxaemia (yes, no), neonatal jaundice (yes, no), twin membership (singleton, monozygotic, and dizygotic, as categorical indicator variables), and birth weight (<2,500, 2,500-2,999, 3,000-3,499, 3,500-3,999, and >=4,000g, as categorical indicator variables).

We found no evidence that neonates who did not conform to the usual growth pattern are at different breast cancer risk when compared to the reference category of neonates who lost less than 200 g after birth and grew at a rate less than 25 g/day after nadir. In contrast, however, neonates who lost ≥200 g after birth, or neonates who grew at a rate of ≥25 g/day after nadir, or neonates with both of these growth pattern characteristics were at an approximately 50% increased risk for breast cancer in later life, in comparison to the reference category. The excess risk was evident and statistically significant among women younger than 50 years old, who were presumably premenopausal at the time of diagnosis of breast cancer (Table CP1.1).

Table CP1.1. Conditional logistic regression-derived* odds ratios (ORs) and 95% confidence intervals (CIs) for breast cancer in relation to patterns of postnatal weight change

	All women			Women <50 years old (presumably premenopausal)		
	OR	95% CI	P-value	OR	95% CI	P-value
Maximum weight loss / daily weight gain since nadir						
<200g / <25g/day	Reference			Reference		
≥200g / <25g/day	1.53	0.96-2.44	0.08	1.81	1.00-3.25	0.048
<200g / ≥25g/day	1.39	0.78-2.45	0.26	2.33	1.13-4.78	0.02
≥200g / ≥25g/day	1.57	0.98-2.51	0.06	2.04	1.12-3.74	0.02
Irregulars (hospitalized for >21 days)	1.37	0.64-2.90	0.42	1.09	0.40-2.95	0.87

* Controlling for maternal age, maternal socioeconomic status, maternal parity, pregnancy toxaemia, neonatal jaundice, twin membership and birth weight

We interpret our findings as indicating that higher levels of pregnancy hormones, several of which have mammotropic properties, and higher levels of growth hormones during the immediate postnatal period, particularly IGF-1, might play an important role in breast cancer, risk, particularly premenopausal breast cancer risk, several decades later.

The findings of this study are intriguing and the apparent magnitude of effect (two-fold increases in premenopausal breast cancer risk for essentially dichotomous contrasts) indicates that the phenomenon is of considerable biomedical importance. Though confidence limits were wide, these findings indicate that the perinatal - including the immediate postnatal - period is likely to play a key role in the natural history of breast cancer and provide support to the hypothesis that early life growth is intimately linked to the risk of this disease.

CP1 reportable outcomes

- Results on the association of postnatal growth with breast cancer risk were presented in the February 2009, 2010 and 2011 LINKS meetings.
- A manuscript on the association of postnatal growth with breast cancer risk has been published (Lagiou et al, 2008 – see Appendix 2).

CP1 Conclusion

The work under this component project was successfully completed. A manuscript has been published.

CP1 References

Lagiou P, Hsieh CC, Trichopoulos D, Adami HO, Hall P, Chie L, Ekbom A. Neonatal growth and breast cancer risk in adulthood. Br J Cancer. 2008;99:1544-8.

CP2: "Relation of perinatal characteristics and postnatal growth velocity with mammographic patterns in adult life"

CP2 PI: Prof. Per Hall, Dept. of Medical Epidemiology and Biostatistics, Karolinska Institutet, P.O. Box 281, SE-171 77 Stockholm, Sweden

Timetable of research accomplishments of CP2 as outlined in the Statement of Work.

Task 2 To investigate the relation of perinatal characteristics and postnatal growth velocity with mammographic patterns in adult life:

a. Retrieval of available birth records from 3,345 women with invasive breast cancer and 3,454 controls (these women are not the same with those to be studied in the context of task 1). (Months 1-48)
b. Retrieval of the available sequential mammographies of the women with breast cancer and the control women. (Months 1-48)
c. Linkage of mammographic data to perinatal characteristics and postnatal growth velocity. (Months 30-48)

d. Evaluation of mammographies through a computer-assisted grey-scale thresholding methods technique. (Months 25-36)
e. Data analyses. (Months 37-48)
f. Manuscript preparation and submission. (Months 49-60)

CP2 progress report

The retrieval of birth records (task 2a) ended in the fall of 2007. The retrieval of available mammograms (task 2b) was a multi-step procedure. In a first step, images were identified and brought to the Department of Medical Epidemiology and Biostatistics at Karolinska. We identified 45,000 images. Due to the existence of multiple images per woman at various examinations, a subset of images per woman and mammographic examination was chosen. We examined the concordance between left and right breast and the two major projections used in mammograms, cranio-caudal (CC) and medio-lateral-oblique (MLO), in a subset of women. There was a very high concordance for both right / left breast and MLO / CC mammographic projections. In Sweden, in contrast to some other countries, nearly all screening examinations are done using the MLO view.

In a second step, all eligible images were digitized. This step was completed in November 2008. We have been able to retrieve and digitize mammograms of the MLO projection for more than 3500 study women. The total number of images digitized exceeds 18000, since we retrieved mammograms from all examinations until date of cancer diagnosis for cases or until the year 1995 for controls, i.e. multiple examinations for each individual.

In a third step, the density of each mammogram was measured. The evaluation of mammograms through a computer-assisted grey-scale thresholding methods technique (task 2d) was done with the Cumulus software, a semi-automatic computer assisted technique where the evaluator indicates the area to be measured (Br J Cancer. 2002;87:876-82). The program was installed, and the responsible Swedish scientists attended a course on the method given by the developers of the program, Norman Boyd and Martin Yaffe in Toronto in April 2007 and 2008 respectively. We chose to measure the last mammogram before the diagnosis of breast cancer for cases and the corresponding date for controls. For cases the opposite breast was chosen. For controls we randomly chose any of the two breasts. This part of the project was finalized in late 2009.

We collaborated with radiologists at Karolinska Hospital, with Drs Isabel dos Santos Silva and Julian Peto at London School of Hygiene and Tropical Medicine, as well as with Dr. Norman Boyd at the Ontario Cancer Institute, Toronto, to assure good quality of the digitizing process and density measurements. Both the London and Toronto group have for many years, studied the relation of mammographic density with breast cancer risk, and are highly experienced in digitizing and evaluating density.

We have conducted analyses on the association of birth characteristics retrieved from birth records (task 2c) with mammographic density in the subsample of women without breast cancer for whom mammographic density was determined and perinatal information from birth records was available (918 women).

A paper reporting a positive association of birth weight with mammographic density among postmenopausal women was published (Tamimi et al, 2010a – see Appendix 2).

In summary:

- All eligible and identified mammograms and birth characteristics were retrieved.
- All retrieved mammograms were digitized, that is, more than 18000 mammograms of the medio-lateral-oblique projection. These digitized mammograms refer to more than 3500 study women.
- Measurements of mammographic density from our group were validated against those of an experienced British group, assuring an overall high quality of the measurements.
- Measurements of mammographic density in all 3,500 women for which mammograms were digitized were completed.
- A paper on the influence of birth characteristics and mammographic density has been published (Tamimi et al, 2010a – see Appendix 2).
- Further exploratory analyses are considered, because there is evidence that density change over time is a better correlate of breast cancer risk than a single mammogram.

CP2 key research accomplishments

Summary: We found that birth size is associated with postmenopausal mammographic density measured more than five decades later. These results support the hypothesis that adult breast density, a powerful correlate of breast cancer risk, has intrauterine roots, as reflected in birth size.

Specifically:
We have studied the association of birth size measurements with mammographic density, a marker of mammary gland mass. For a population-based sample of 893 postmenopausal women without previous cancer in Sweden, we retrieved information on birth size from birth records and their most recent mammography. Film mammograms of the medio-lateral oblique view were digitized and the Cumulus software was used for computer-assisted semi-automated thresholding of mammographic density.

We used generalized linear models adjusted for covariates to determine the mean percentage mammographic density without transformation, according to birth size categories. To determine whether there was a linear trend of mean percent breast density with increasing birth size, we calculated p values for inclusion of birth size as an ordered (per category) or continuous (per standard deviation—SD) variable in the model. We included in the multivariate models, as possible confounders, the following known predictors of mammographic density: age at mammogram, body mass index (BMI), parity, and age at menopause. Odds ratios (OR) and 95% confidence intervals (CI) for having high versus low mammographic density also were determined using multiple logistic regression. To determine whether there was a linear trend with increasing birth size, we calculated p values from Wald statistics including a continuous

term (per SD) in the model. All p values were derived from two-sided tests of statistical significance.

Table CP2.1: Mean percent mammographic density (MD) according to birth size characteristics

	N	Mean MD (%)[1]	Mean MD (%)[2]	Mean MD (%)[3]
Birth weight (g)				
≤2500	32	14.9	15.6	15.6
2501-≤3000	145	13.8	13.5	13.6
3001-≤3500	324	15.6	15.8	15.8
3501-≤4000	291	18.1	17.8	17.7
>4000	101	17.6	18.3	18.6
P-trend[4]		0.04	0.02	0.02
Birth length (cm)				
<50	232	16.2	16.3	16.4
50	204	16.3	16.1	16.2
51	140	15.6	15.6	15.5
52	130	18.3	18.0	18.1
53+	156	16.5	16.7	16.5
P-trend[4]		0.71	0.83	0.69
Head circumference (cm)				
<34	170	15.8	16.1	15.5
34-<35	197	16.6	16.7	16.9
35-<36	192	16.3	16.1	16.2
36+	170	20.4	20.2	20.4
P-trend[4]		0.04	0.03	0.007

[1] Adjusted for age (continuous)
[2] Adjusted for age (continuous) and body mass index (continuous)
[3] Adjusted for age (continuous), body mass index (continuous), age at menopause (<45, 45-49, 50-54, 54+, missing categorical), parity (0, 1, 2, 3+, categorical)
[4] P for trend based on per standard deviation increase

Table CP2.1 shows mean percent mammographic density according to categories of birth size characteristics. The mean values are adjusted first for age at mammogram, then for age at mammogram and BMI, and then for age at mammogram, BMI, age at menopause and parity. Mean percent mammographic density was significantly positively associated with both birth weight and head circumference. In the full models, p values for trend were 0.02 and 0.007, respectively. Thus, women with birth weight exceeding 4000 grams had a mean percent mammographic density of 18.6%, while those with birth weight 2500 grams or less had a mean percent mammographic density of 15.6%. Moreover, women with a birth head circumference of 36 cm or more had a mean percent mammographic density of 20.4%, while those with a circumference less than 34 cm had a mean percent mammographic density of 15.5%.

Both of these associations were driven by the positive association with dense area on the mammogram (p for trend 0.02 for birth weight and 0.01 for head circumference), as there was no association with non-dense area on the mammogram (p for trend 0.15 for birth weight and 0.16 for head circumference) in multivariate models.

Table CP2.2. Odds ratios (OR) and 95% confidence intervals (95% CI) for high versus low mammographic density in relation to birth size characteristics, using 50% density as cutoff.

	Mammographic density		
Birth weight (g)	High (≥50%)	Low (<50%)	OR (95%CI)[1]
≤2500	3	29	0.57 (0.18-1.81)
2501-≤3000	2	143	
3001-≤3500	13	311	1.0 (REF)
3501-≤4000	21	270	1.68 (0.78-3.62)
>4000	8	93	2.91 (1.07-7.88)
P-trend[2]			0.048
Birth length (cm)			
<50	11	221	0.77 (0.27-2.21)
50	9	195	0.85 (0.29-2.52)
51	7	133	1.0 (REF)
52	9	121	1.10 (0.37-3.26)
53+	10	146	1.04 (0.35-3.11)
P-trend[2]			0.49
Head circumference (cm)			
<34cm	10	160	0.66 (0.23-1.87)
34-<35cm	9	188	0.90 (0.33-2.44)
35-<36cm	10	182	1.0 (REF)
36+	15	155	1.72 (0.68-4.35)
P-trend[2]			0.04

[1] Adjusted for age (continuous), body mass index (continuous), age at menopause (<45, 45-49, 50-54, 54+, categorical), parity (0, 1, 2, 3+, categorical)
[2] P for trend based on per standard deviation increase

To further explore the pattern of association of birth size characteristics with mammographic density, we have calculated odds ratios for high versus low mammographic density in relation to birth size characteristics (Table CP2.2). For all three birth size characteristics, intermediate categories were used as referents. In line with the results shown in Table CP2.1, birth weight and head circumference were positively associated with mammographic density. The trends were monotonic and statistically significant only when the cutoff was set at 50% (and not at 25%), indicating that the association between birth size and mammographic density is not linear and is particularly evident at the extreme. Again, in line with the results in Table CP2.1, we found no association of birth length with mammographic density.

We have also explored whether the association between birth size and mammographic density might be mediated through measures of growth and development, notably age at menarche and adult height, by alternatively and simultaneously including these variables in the models. Height and, to a lesser extent, age at menarche appear to mediate, in part, the associations of both birth weight and head circumference with mammographic density.

The results from this large population-based sample of postmenopausal Swedish women, indicate that birth weight and head circumference are both significantly positively associated with postmenopausal mammographic density. The 3–5% difference in mammographic density

comparing the extreme categories of birth weight or head circumference is not trivial and is indeed comparable to that observed between women with and without hormonal treatment. Although birth length was not associated with mammographic density, this could reflect the fact that birth length is measured less reliably than birth weight and head circumference. Moreover, we found evidence that part of the association between birth size and mammographic density could be explained by the association of birth size with height and, to a lesser extent, with age at menarche.

This is the first study to evaluate birth size parameters retrieved from birth records in relation to mammographic density assessed by a semi-automated method. Earlier studies assessed breast density using a more subjective categorical measure of breast density or relied on self-reported birth weight. Computer-assisted thresholding methods, as those used in our study, which are less subjective than categorical classifications, have been demonstrated to be better predictors of breast cancer risk. In addition, birth data from records are more reliable than those relying on self-recall.

Mammographic density is one of the strongest known risk factors for breast cancer, with women in the highest category of mammographic density being at a four- to sixfold increased risk of breast cancer as compared with women in the lowest category. Although the biologic mechanism underlying this risk remains unclear, it is possible that mammographic density is influenced in part by exposures early in life, including the intrauterine life. In the context of the present Innovator award, it has been shown that birth weight is positively associated with the size of the cord blood stem cell pool, a surrogate for overall stem cell potential (Br J Cancer. 2008;98:660-3). Thus, assessed through mammographic density, mammary gland mass is likely to be correlated with the number of mammary stem cells, generating the association of birth weight with mammographic density.

In conclusion, our results support the hypothesis that adult breast density, a powerful correlate of breast cancer risk, has intrauterine roots, as reflected in birth size. The results indirectly provide further support to the general hypothesis that breast cancer risk in adulthood is programmed already in utero and then modified by a multitude of factors later in life.

CP2 reportable outcomes

- Results on the association of birth size with postmenopausal mammographic density were presented in the February 2009, 2010 and 2011 LINKS meetings.
- A manuscript on the association of birth size with postmenopausal mammographic density has been published (Tamimi et al, 2010a – see Appendix 2).

CP2 Conclusion

As indicated in the previous progress reports, the rate of retrieving information on birth characteristics and mammographic images was in line with what we anticipated and was finished in the fall of 2007 and 2008 respectively. The digitizing rate was lower than initially anticipated, which is why an additional person was hired in 2008 to speed up the process. The rate of density

measurements was also slower than originally anticipated due to the large amount of mammograms. Nevertheless, these obstacles were overcome. The evaluation of mammographies through a computer-assisted grey-scale thresholding methods technique was successfully completed, data analyses were undertaken and a manuscript was published. The work under this component project was successfully completed. Further exploratory analyses are considered.

CP2 References

Tamimi RM, Eriksson L, Lagiou P, Czene K, Ekbom A, Hsieh CC, Adami HO, Trichopoulos D, Hall P. Birth weight and mammographic density among postmenopausal women in Sweden. Int J Cancer. 2010a;126:985-91.

CP3: "Interaction of perinatal characteristics with genes that are likely related to breast cancer risk"

CP3 PI: Prof. Per Hall, Dept. of Medical Epidemiology and Biostatistics, Karolinska Institutet, P.O. Box 281, SE-171 77 Stockholm, Sweden

Timetable of research accomplishments of CP3 as outlined in the Statement of Work.

Task 3 To investigate the possible interaction of perinatal characteristics with genes that are likely related to breast cancer risk:

a. Identification and selection of genes likely to be related to breast cancer risk, e.g. ESR1, AIB1, and the IGF family (Months 1-12)
b. Selection of "tagging" single nucleotiod polymorphisms (tSNPs). The choice of tSNBs aims at avoiding redundant genotyping. A good marker coverage is expected to be achieved by using approximately one SNP per 3 Kb. (Months 1-12).
c. Genotyping of the approximately 8 genes selected for the study (Months 13-36).
d. Data analyses. (Months 36-48)
e. Manuscript preparation and submission. (Months 49-60)

CP3 progress report

Identification and selection of genes likely to be related to breast cancer risk, (Task 3a), selection of "tagging" single nucleotiod polymorphisms (tSNPs) (Task 3b) and genotyping in the selected genes (Task 3c) were completed. Data analyses (Task 3d) and manuscript preparation (Task 3e) were also completed and a manuscript has been published (Tamimi et al, 2010b – see Appendix 2).

In addition to the published results, we also contributed to a study evaluating genetic variation in the estrogen metabolic pathway in relation to mammographic density (Breast Cancer Research

2010;12:R19– with acknowledgment to support by W81XWH-05-1-0314 Innovator Award) with a view to eventually study possible interactions with perinatal characteristics (see Appendix 4).

In summary:

- Associations of loci with breast cancer risk were evaluated in our data and were found compatible with what was reported from previous GWAS studies.
- Interaction of birth weight with higher breast cancer risk SNPs in relation to breast cancer occurrence was evaluated.
- A paper was published (Tamimi et al, 2010b – see Appendix 2).

CP3 key research accomplishments

Summary: We found suggestive evidence for interaction of birth weight with genetic predisposition in relation to breast cancer risk, notably with respect to SNP rs2981582. Hence, the size of the pool of mammary stem cells, as reflected in birth weight, appears to interact with genetic susceptibility in modulating breast cancer risk.

Specifically:
There is considerable evidence that birth weight is positively associated with breast cancer risk, and seven single-nucleotide polymorphisms (SNPs) have been conclusively associated with this risk. We have hypothesized that breast cancer susceptibility loci may have a greater influence on breast cancer risk among women with higher birth weight, who are expected to have a larger pool of mammary stem cells that are susceptible to malignant transformation.

In light of what has been published in the international literature with respect to important polymorphisms associated with breast cancer risk (Nature. 2007;447:1087-93; Breast Cancer Res. 2007;9(6):R78; Nat Genet. 2007;39:865-9; Nat Genet. 2007;39:870-4; Breast Cancer Res. 2008;10(4):R66), we looked at these polymorphisms in relation to breast cancer risk in our data. We studied the following single nucleotide polymorphisms (SNPs): rs2981582 (in the FGFR2 gene that encodes the fibroblast growth factor receptor 2), rs12443621 (in the trinucleotide repeat containing 9, TNRc9 gene), rs8051542 (in the TNRc9 gene), rs889312 (in the MAP3K1 gene that encodes mitogen-activated protein kinase kinase kinase 1), rs3817198 (in the LSP1 gene that codes for lymphocyte-specific protein 1), rs13281615 (on 8q24), rs3803662 (in the TNRc9 gene).

In the context of a nationwide, population-based case–control study in Sweden, genotype analysis was performed for 1,314 women with breast cancer and 1,515 control women. We isolated DNA from 3 ml of whole blood with the Wizard Genomic DNA Purification Kit (Promega, Madison, WI) in accordance with the manufacturer's instructions. Genotyping of the selected SNPs was conducted by Sequenom iPLEX and Taqman.

We attempted to identify birth records for breast cancer cases and controls, using the Swedish national registration numbers. Using information from the birth registry, research assistants visited 120 hospital and medical record archives to locate the original birth records for the study

participants. Because this study included births occurring between 1918 and 1945 throughout all of Sweden, a large number of births at this time had occurred at home, and birth characteristic information had never been measured or recorded. We were able to locate birth records or midwife journals with adequate information on birth characteristics for about 50% of the participants. Using a detailed form, perinatal characteristic information from available birth records was abstracted. Birth weight in grams and twin status were consistently listed in birth records. Less consistently, information was provided on birth length, head circumference and placental weight. In total, there were 710 breast cancer cases and 770 controls with both genotype information and birth weight data. We excluded cases and controls who were twin members and/or had missing information on one or more covariates. Eventually, we studied 693 women with breast cancer and 747 control women, for whom both birth characteristic data and information on at least one of the evaluated SNPs were available.

Initially, we examined the association between breast cancer risk and each of the following selected SNPs: rs2981582 (FGFR2), rs3803662 (TNRC9/LOC643714), rs889312(MAP3K1), rs13281615 (8q24), rs3817198 (LSP1), rs12443621 (TNRC9/LOC643714), and rs8051542 (TNRC9/LOC643714). For each SNP, odds ratios (OR) and 95% confidence intervals (CI) for breast cancer risk were determined using logistic regression. To determine whether there was a linear trend with increasing number of variant alleles in each SNP, we calculated p values from Wald statistics including a continuous term in the model. Tests for interaction were based on a Wald test after inclusion of an interaction term. All p values presented are from two-sided tests of statistical significance (Table CP3.1).

Three of the breast cancer susceptibility loci identified from the BCAC genome-wide association study (Nature. 2007;447:1087-93) were significantly associated with breast cancer risk in this population with the strongest association for rs2981582, in line with previous reports. Women homozygous for this allele had an almost twofold (OR = 1.99, 95% CI 1. 45–2.72) increased risk of breast cancer relative to women homozygous for the wild-type allele, whereas women heterozygous for this SNP were at higher risk of breast cancer compared to homozygotes for the wild-type allele (OR = 1.25, 95% CI 0.99–1.57). Of the remaining polymorphisms, trend tests with increasing number of high-risk alleles were statistically significant in two instances (rs12443621 and rs3803662).

Because rs2981582 in intron 2 of FGFR2 has been repeatedly identified as the top hit in GWAS of breast cancer (Nature. 2007;447:1087-93; Nat Genet. 2007;39:870-4) and has demonstrated the largest excess risk in our study, we first examined the interaction between rs2981582 and birth weight in relation to breast cancer risk (Table CP3.1). To increase power without invoking linear trends, we grouped heterozygotes with homozygotes for the wild-type allele, because the difference in risk between homozygotes for the variant allele and heterozygotes is substantially larger than that between heterozygotes and homozygotes for the wild-type allele. We found a marginally significant interaction between birth weight and rs2981582 (p for interaction 0.07). As shown in Table CP3.1, the interaction is generated by a strong, though marginally significant, positive association between birth weight and breast cancer risk among homozygotes for the variant allele (p for trend 0.07) and a weak non-significant inverse association between birth weight and breast cancer risk among heterozygotes and homozygotes for the wild-type allele (p

for trend 0.68). Among women with the highest category of birth weight (>4,000 g), being homozygous for the rs2981582 variant was associated with an over fivefold increased risk of breast cancer (OR = 5.52, 95% CI 1.96–15.56) relative to women with wild-type alleles and similar birth weight.

We also examined a potential interaction between birth weight and the other 6 SNPs (Table CP3.1), following the same approach used for rs2981582. None of the other SNPs demonstrated a significant interaction with birth weight, but we generally observed a positive association between birth weight and breast cancer risk among women homozygous for the high-risk allele.

Table CP3.1: Odds ratios (OR) and 95% confidence intervals (95% CI) for breast cancer according to birth weight and susceptibility SNP genotype

rs2981582	**WT/WT and WT/VAR**	**VAR/VAR**
Birth weight (g)	**OR(95%CI)**	**OR(95%CI)**
	Cases/controls	**Cases/controls**
≤2500	1.21 (0.64-2.29)	0.86 (0.30-2.53)
	21/20	6/8
2501-≤3000	0.89 (0.63-1.26)	2.15 (1.12-4.13)
	80/103	28/15
3001-≤3500	1.0 (REF)	1.27 (0.80-2.02)
	211/243	44/40
3501-≤4000	1.02 (0.77 -1.34)	1.90 (1.10-3.30)
	174/197	38/23
>4000	0.84 (0.57-1.23)	4.61 (1.70-12.49)
	58/80	20/5
P-value for trend	0.68 (negative trend)	0.07 (positive trend)
P-value for interaction	0.07	
rs12443621		
Birth weight (g)		
≤2500	1.36 (0.73-2.51)	0.87 (0.24-3.11)
	23/22	4/6
2501-≤3000	1.24 (0.88-1.76)	0.92 (0.48-1.76)
	90/94	17/24
3001-≤3500	1.0 (REF)	2.03 (1.34-3.08)
	185/240	72/46
3501-≤4000	1.28 (0.96-1.70)	1.21 (0.76-1.92)
	171/174	42/45
>4000	1.03 (0.70-1.51)	2.31 (1.00-5.34)
	61/77	16/9
P-value for trend	0.78 (negative trend)	0.30 (positive trend)
P-value for interaction	0.29	

(continued)

Table CP3.1 (continued): Odds ratios (OR) and 95% confidence intervals (95% CI) for breast cancer according to birth weight and susceptibility SNP genotype

rs8051542		
Birth weight (g)		
≤2500	1.63 (0.86-3.09)	0.41 (0.11-1.53)
	24/18	3/9
2501-≤3000	1.14 (0.81-1.60)	1.09 (0.55-2.16)

	92/99		17/19
3001-≤3500	1.0 (REF)		1.66 (1.10-2.51)
	190/232		68/50
3501-≤4000	1.26 (0.95-1.66)		0.80 (0.48-1.33)
	184/179		28/43
>4000	1.07 (0.72-1.58)		1.40 (0.66-2.93)
	63/72		16/14
P-value for trend	0.92 (negative trend)		0.81 (positive trend)
P-value for interaction		0.80	
rs889312			
Birth weight (g)			
≤2500	1.15 (0.65-2.03)		--
	26/26		0/3
2501-≤3000	1.09 (0.79-1.50)		0.69 (0.25-1.93)
	102/108		6/10
3001-≤3500	1.0 (REF)		1.36 (0.75-2.46)
	229/263		26/22
3501-≤4000	1.08 (0.83-1.41)		1.68 (0.81-3.47)
	194/206		19/13
>4000	1.03 (0.72-1.49)		1.15 (0.37-3.61)
	72/80		6/6
P-value for trend	0.91 (negative trend)		0.13 (positive trend)
P-value for interaction		0.14	
rs3817198			
Birth weight (g)			
≤2500	1.05 (0.58-1.89)		1.14 (0.28-4.59)
	23/25		4/4
2501-≤3000	1.04 (0.76-1.44)		1.01 (0.38-2.66)
	100/109		8/9
3001-≤3500	1.0 (REF)		1.08 (0.55-2.10)
	235/267		18/19
3501-≤4000	1.15 (0.88-1.50)		0.89 (0.47-1.69)
	197/195		18/23
>4000	0.89 (0.62-1.29)		4.55 (1.27-16.30)
	65/83		12/3
P-value for trend	0.92 (negative trend)		0.26 (positive trend)
P-value for interaction		0.27	

(continued)

Table CP3.1 (concluded): Odds ratios (OR) and 95% confidence intervals (95% CI) for breast cancer according to birth weight and susceptibility SNP genotype

rs13281615			
Birth weight (g)			
≤2500	1.50 (0.72-3.13)		0.55 (0.22-1.37)
	18/14		7/15
2501-≤3000	1.26 (0.86-1.86)		0.72 (0.42-1.24)
	79/73		26/42
3001-≤3500	1.0 (REF)		1.16 (0.81-1.65)

	143/167		103/104
3501-≤4000	1.29 (0.93-1.78)		0.89 (0.60-1.32)
	143/130		63/83
>4000	1.19 (0.77-1.84)		0.97 (0.54-1.74)
	55/54		24/29
P-value for trend	0.97 (positive trend)		0.46 (positive trend)
P-value for interaction		0.56	
rs3803662			
Birth weight (g)			
≤2500	1.15 (0.64-2.06)		0.29 (0.03-2.60)
	25/25		1/4
2501-≤3000	1.05 (0.76-1.45)		1.30 (0.49-3.42)
	100/110		9/8
3001-≤3500	1.0 (REF)		1.65 (0.92-2.95)
	230/265		30/21
3501-≤4000	1.13 (0.87-1.47)		0.89 (0.38-2.06)
	202/206		10/13
>4000	1.07 (0.75-1.53)		1.15 (0.29-4.66)
	76/82		4/4
P-value for trend	0.76 (positive trend)		0.87 (positive trend)
P-value for interaction		0.94	

WT: wild type
VAR: high risk variant

It has been suggested that the documented association of breast cancer risk with birth weight could reflect the underlying association of this risk with the size of the mammary stem cell pool (Br J Cancer 2008;98:660–663) and eventually mammary gland mass (Mutat Res 1995;333:29–35; Breast Cancer Res 2005;7:13–17). In women who carry high risk alleles, the association of birth weight (as a correlate of mammary stem cell pool) with breast cancer risk would be expected to be stronger. Our findings are compatible with this hypothesis, although the study was not sufficiently powered to document interactions of modest strength. Of note, our study did not document the main effect of four of the seven SNPs identified in the context of GWAS (Nature. 2007;447:1087-93). When our study was initiated, the effect size of SNPs that could be related to breast cancer was not known, precluding reliable power calculations. The evidence that the association between birth weight and breast cancer could be modified by genetic susceptibility was stronger for the SNP in FGFR2 than for any of the other susceptibility loci. SNPs in intron 2 of FGFR2 have emerged as top hits from multiple GWAS of breast cancer (Nature. 2007;447:1087-93; Nat Genet. 2007;39:870-4) and have been significantly associated with breast cancer risk in a number of populations including European, Asian, Ashkenazi Jewish, and African American women. FGFR2, a tyrosine kinase receptor belonging to a family of genes involved in growth and proliferation, is overexpressed in breast tumors and may function as an oncogene. An interpretation of our results is that susceptibility loci in an oncogene such as FGFR2 put individual cells at higher risk of malignant transformation. To the extent that birth weight is a proxy for glandular mass, having an increased number of mammary stem cells with high-risk alleles would be associated with a disproportionally increased risk of breast cancer.

A possible interaction between genetic susceptibility loci and birth weight in relation to breast cancer risk has not been previously investigated. In fact, there are few reliable studies of gene–environment interactions and breast cancer risk. Our results are compatible with the hypothesis that the pool of mammary stem cells is critical in the intrauterine roots of breast cancer risk in adult life. This is because the size of the pool of mammary stem cells, as reflected in birth weight, appears to interact with genetic susceptibility in modulating breast cancer risk.

CP3 reportable outcomes

- Results on the interaction of birth weight with higher breast cancer risk SNPs in relation to breast cancer occurrence were presented in the February 2010 and 2011 LINKS meetings.
- A manuscript on the interaction of birth weight with higher breast cancer risk SNPs in relation to breast cancer has been published (Tamimi et al, 2010b – see Appendix 2).

CP3 Conclusion

The work under this component project was successfully completed. A manuscript has been published and further exploratory analyses are considered.

CP3 References

Tamimi RM, Lagiou P, Czene K, Liu J, Ekbom A, Hsieh CC, Adami HO, Trichopoulos D, Hall P. Birth weight, breast cancer susceptibility loci, and breast cancer risk. Cancer Causes Control. 2010b;21:689-96.

CP4: "Pregnancy hormones and perinatal breast cancer risk factors in Boston, USA and Shanghai, China"

CP4 co-PIs: Prof. Dimitrios Trichopoulos and Dr. Pagona Lagiou, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115,

Timetable of research accomplishments of CP4 as outlined in the Statement of Work.

Task 4 To study maternal and cord blood levels of components of the IGF system and adiponectin among Caucasian women in North America and Chinese women in Asia in conjunction to maternal anthropometry and birth size parameters:

a. Retrieval of the available stored cord blood and maternal serum samples from 304 pregnant Caucasian women in Boston, US and 335 pregnant Chinese women in Shanghai, China and transfer of these samples to the laboratory for hormone determination. (Months 1-6)

b. Conduct of laboratory assays for hormones (Months 7-24)
c. Linkage of maternal and newborn data to maternal and cord blood hormone levels (Months 24-30)
d. Data analyses (each of the measured hormones in maternal and cord blood to be studied in conjunction to maternal and newborn variables) (Months 31-48)
e. Manuscript preparation and submission. (Months 37-60)

CP4 progress report

The database regarding baseline data was already in place by the first annual report.

During the second reporting period, hormone determinations by the UMass Medical School, ILAT Steroid RIA Laboratory were completed for 661 serum samples and 114 cord blood samples provided by pregnant women recruited in Boston, as well as for 762 serum samples and 131 cord blood samples provided by pregnant women recruited in Shanghai, China. Testosterone, IGF-1, IGFBP-3, and adiponectin were measured on the maternal samples, while for cord blood samples, in addition to these hormones, estradiol, estriol, progesterone, SHBG and IGF-2 were measured (measurements of these hormones in the maternal sera had already been conducted in the context of an earlier NIH study).

During the third reporting period, data on hormone determinations in the serum and cord blood samples were merged with the baseline information on pregnant women and their newborn. The updated database contains information on: (i) baseline data and serum hormone measurements from 302 eligible women in Boston and 339 eligible women in Shanghai and (ii) baseline data and cord blood hormone measurements from 111 eligible newborn in Boston and 122 eligible newborn in Shanghai. Exploratory analyses were conducted and two papers were published (Cancer Epidemiol Biomarkers Prev. 2007;16:1720-3; Cancer Epidemiol Biomarkers Prev. 2008;17:224-31). In these previous analyses, about half the samples from Boston and less than a quarter of the samples from Shanghai were analyzed with respect to IGF-1 and IGFBP-3.

In the fourth reporting period, measurements of cord blood levels of IGF-1, IGF-2 and IGFBP-3 in all available samples from Boston, US and Shanghai, China were undertaken for the first time, using state-of-the-art analytical assays.

During the fifth reporting period a manuscript was published (Lagiou et al, 2009 – see Appendix 2).
During the one year extension granted to the Innovator project, two additional manuscripts were published. One referred to the core aims of the Innovator (Lagiou et al, 2010 – see Appendix 2) and the second to a peripheral issue (with explicit reference to support by the Innovator project) (see Appendix 4). Further analyses are planned to evaluate the potential effect of maternal diet during pregnancy on the levels of IGF components in maternal sera and cord blood.

CP4 key research accomplishments

Summary:

- Cord blood levels of IGF-1 are substantially higher among Caucasian babies born in Boston compared to Chinese in Shanghai. The documentation of higher cord blood levels of IGF-1 - a principal growth hormone which does not cross the placenta - among Caucasian compared to Asian newborns is concordant with breast cancer incidence in these populations.
- Cord blood levels of IGF-2 are substantially higher among Chinese babies born in Shanghai compared to Caucasian babies born in Boston.
- Cord blood IGF-1 is significantly positively associated with birth weight and birth length in Boston, but not Shanghai.
- Positive, though statistically non-significant, associations of cord blood IGF-2 with birth size are only evident in Shanghai.
- For cord blood IGF-1 in Boston and for cord blood IGF-2 in Shanghai, the associations with birth weight and birth length are positive and significant among taller women (among whom maternal anthropometry is likely to impose fewer constrains) but essentially null among shorter women.
- Although maternal hormone levels at the second and third trimester of pregnancy are highly intercorrelated, they are essentially unrelated to core blood levels of the same compounds.
- Both estradiol and testosterone are significantly higher in the cord blood among Chinese un Shanghai compared to Caucasians in Boston, but it is cord blood levels of SHBG which are strikingly higher among the former compared to the latter. These results point to SHBG as an important modulator of early life influences on adult life breast cancer risk.

Specifically:

In the context of CP4 a manuscript focusing on the role of cord blood IGF compounds was published in 2009 (Lagiou et al, 2009 – see Appendix 2).

Breast cancer incidence and birth weight are higher among Caucasian than Asian women (Int J Cancer 2001;94:153-6), and birth size has been positively associated with breast cancer risk (PLoS Med 2008; 5: e193). Maternal pregnancy hormone levels, however, notably estrogens and androgens, have been generally lower in Caucasian than Asian women (Brit J Cancer 1999;79:7-12). We hypothesized that cord blood levels of insulin-like growth factor-1 IGF-1, which does not cross the placenta (Placenta 1999;20:325–330), are higher among Caucasian than among Asian neonates, and that cord blood IGF-1 levels are positively associated with birth size. We evaluated this hypothesis, and also examined the role of IGF-2, a main component of the IGF system in fetal life, by studying cord blood samples from babies born to women in Boston, USA and Shanghai, China.

We were able to collect cord blood samples of acceptable quality and sufficient quantities for the determinations of cord blood hormones for 202 uncomplicated full-term pregnancies (37–42 weeks long without pregnancy toxemia), 92 in Boston and 110 in Shanghai. Hormone determinations were conducted at the ILAT Steroid RIA Laboratory of the University of Massachusetts Medical School. IGF-1, IGF-2, and IGFBP-3 were measured by coated-tube immunoradiometric assay kits (Diagnostic System Laboratories, Inc., Webster, TX, USA). The laboratory-estimated inter-assay and intra-assay coefficients of variation were, respectively,

9.0% and 3.3% for IGF-1, 5.9% and 3.4% for IGF-2, and 8.0% and 4.8% for IGFBP-3. There was no detectable cross reactivity of the IGF-1 assay with IGF-2 according to the manufacturer's specificity assessment.

Multiple regression models were used to compare hormone levels between Boston and Shanghai controlling for maternal age, height, duration of gestation and weight gain (all continuously), as well as for parity and gender of offspring. The association of IGF-1 and IGF-2 with birth weight and birth length were examined by modeling the data through multiple regression models with, alternatively, birth weight and length as the outcomes, controlling also mutually for IGF-1 and IGF-2, as well as for IGFBP-3. Analyses were conducted separately for each centre, for all women, as well as for tall and short women, with the cut-off for height set a priori at 163 cm (median height in Boston, third quartile in Shanghai). The hypothesis is that, among shorter women, maternal anthropometry imposes constraints on birth size and could trigger negative feedback mechanisms that could obscure positive associations of hormones with birth size.

In Table CP4.1, cord blood levels of IGF-1, IGF-2, and IGFBP-3 are compared between centers, adjusting for maternal age, height, weight gain, parity, duration of gestation, and gender of offspring. Levels of IGF-1 were significantly and substantially higher in Boston, whereas IGF-2 levels were significantly lower. The differences in IGF-1 and IGF-2 were amplified after adjustment both mutually as well as for IGFBP-3.

Table CP4.2 shows multiple regression-derived partial regression coefficients of birth weight (upper panel) and birth length (lower panel) on one standard deviation increments of cord blood levels of IGF-1 and IGF-2 in Boston and Shanghai, both overall and in strata defined by maternal height. In Boston, with respect to both birth weight and birth length, there were significant positive associations with IGF-1; stratified analyses indicated that these differences were generated exclusively by the offspring of taller women. In contrast, in Shanghai, no significant associations of IGF-1 with either birth weight or birth length were evident.

The results for IGF-2 were strikingly different. In Boston, no association was evident with respect to either birth weight or birth length, neither among women overall nor among taller or shorter women. In Shanghai, however, cord blood IGF-2 was suggestively positively associated with birth length ($P \sim 0.09$), whereas, among taller women, it was significantly positively associated with both birth weight and length.

Table CP4.1 Percent differences* of cord blood levels of IGF-1, IGF-2 and IGFBP-3 between newborns in Boston, USA (reference) and Shanghai, China

	Unadjusted for the other IGF components		Adjusted for the other IGF components	
	Shanghai vs. Boston (%)	p	Shanghai vs. Boston (%)	p
IGF-1 (ng/ml)	-39.6 (-54.3,-20.2)	0.0005	-50.9 (-61.0,-38.1)	$<10^{-6}$
IGF-2 (ng/ml)	20.7 (7.7,35.2)	0.001	23.2 (10.9,37.0)	0.0001
IGFBP-3 (%)	23.0 (-8.1,64.7)	0.162	22.5 (-5.1,57.9)	0.118

* Adjusted for maternal age, height and weight gain, parity, duration of gestation and gender of offspring. Hormone levels were log-transformed, so that the coefficients express % differences between centers.

Table CP4.2 Multiple regression-derived partial regression coefficients* (and 95% confidence intervals, CIs) of birth weight (upper panel) and birth length (lower panel) on one standard deviation increments of cord blood levels of IGF-1, IGF-2 and IGFBP-3 in Boston, USA and Shanghai, China, overall and by maternal height.

	All Women				Women ≤1.63m height				Women >1.63m height			
	Boston (n=92)		Shanghai (n=110)		Boston (n=46)		Shanghai (n=81)		Boston (n=46)		Shanghai (n=29)	
	b (CI)	p	b (CI)	p	b (CI)	p	b (CI)	p	b (CI)	p	b (CI)	p
BIRTH WEIGHT												
IGF-1 (per 44.4 ng/ml)	141.7 (11.3,272.1)	**0.03**	-3.7 (-133.5,126.0)	0.96	-39.6 (-252.3,173.1)	0.71	-1.2 (-139.9,137.4)	0.99	260.4 (92.6,428.1)	**0.003**	-33.6 (-420.6,353.3)	0.86
IGF-2 (per 132.1 ng/ml)	7.0 (-131.4,145.3)	0.92	57.3 (-51.3,166.0)	0.30	45.5 (-137.2,228.1)	0.62	-2.3 (-124.1,119.5)	0.97	1.0 (-196.7,198.7)	0.99	292.8 (0.59,584.9)	**0.05**
BIRTH LENGTH												
IGF-1 (per 44.4 ng/ml)	0.85 (0.19,1.51)	**0.01**	0.00 (-0.91,0.91)	0.99	0.29 (-0.70,1.28)	0.56	-0.21 (-1.39,0.98)	0.73	1.14 (0.09,2.19)	**0.03**	0.81 (-0.19,1.81)	0.11
IGF-2 (per 132.1 ng/ml)	0.10 (-0.60,0.80)	0.77	0.65 (-0.11,1.41)	0.09	0.44 (-0.41,1.29)	0.30	0.73 (-0.31,1.77)	0.17	-0.17 (-1.41,1.06)	0.78	0.78 (0.02,1.53)	**0.04**

*Adjusted for maternal age, height and weight gain, parity, duration of gestation and gender of offspring in models for all women; adjusted for maternal age and weight gain, and duration of gestation in models by maternal height. In all models, IGF-1 and IGF-2 were adjusted both mutually and for IGFBP-3.

Cord blood IGF-1 has been shown to be positively associated with the size of the stem cell pool (Breast Cancer Res. 2007;9(3):R29), which has also been linked to birth size (Br J Cancer. 2008;98:660-3), a predictor of breast cancer risk in adult life (PLoS Med 2008; 5: e193). To our knowledge, there are no reports concerning a possible association of the stem cell pool with IGF-2, which is a growth promoting hormone during gestation. The differential actions of IGF-1 and IGF-2 in embryonic life could be explained by the fact that both IGFs are known to bind to the signalling IGF-1 receptor, whereas IGF-2 also binds to the nonsignalling IGF-2 receptor (J Clin Endocrinol Metab 2000;85: 4266–4269).

Irrespectively of the underlying physiologic mechanisms, the difference in birth size between Caucasian and Asian newborn can account for only a small fraction of the differences in breast cancer incidence. However, the fact that endocrine perinatal influences on birth size are evident mostly, or exclusively, among newborn of taller women (Table CP4.2) may explain the sharp contrast in breast cancer incidence between Caucasian and Asian women—birth size is known to be positively associated with adult height and, over successive generations, improved nutrition, leading to increased adult body size, might reduce constraints on fetal growth and birth size, which in turn affects adult height. The cycle tends to repeat itself, notably over consecutive generations of Asians migrating to the west, who show a gradual increase of breast cancer incidence, as elaborated in a commentary published in PLoS Medicine (PLoS Med 2008;5: e194) in the context of the present Innovator project (see Appendix 3).

In conclusion, in manuscript Br J Cancer. 2009;100:1794-8 (Lagiou et al, 2009 – see Appendix 2) we reported that cord blood IGF-1 is significantly and substantially higher among Caucasian compared with Chinese babies, whereas the opposite is noted for IGF-2. Moreover, IGF-1 is significantly positively associated with birth weight and birth length in Boston, but not Shanghai. In contrast, positive associations of IGF-2 with birth size are only evident in Shanghai. The positive associations of birth weight and birth length with IGF-1 in Boston and with IGF-2 in Shanghai were significant in taller women, among whom maternal anthropometry does not exercise strong constraints in fetal growth. The documentation of higher cord blood levels of IGF-1, a principal growth hormone that does not cross the placenta, among Caucasian than in Asian newborns is concordant with breast cancer incidence in these populations. On the basis of these results, it appears that cord blood insulin-like growth factors IGF-1 and IGF-2 may have distinct roles for adult life breast cancer risk.

In another manuscript published in 2010 the context of CP4 during the one year no cost extension granted to the project (Lagiou et al, 2010 – see Appendix 2), we studied the 241 singleton pregnancies and offspring of Caucasian women in Boston, USA and the 295 singleton pregnancies and offspring of Chinese women in Shanghai, China and we determined a range of maternal blood hormones at the 16^{th} (samples available from 232 Caucasian and 279 Chinese pregnancies) and 27^{th} (samples available from 225 Caucasian and 281 Chinese pregnancies) gestational week, as well as in umbilical cord blood (samples available from 92 Caucasian and 110 Chinese pregnancies).

Table CP4.3**.** Spearman correlation coefficients between maternal serum levels of the indicated hormones at the 16th and 27th gestational week and levels of these hormones in the cord blood in Boston, USA and Shanghai, China.

	Boston	**Shanghai**
Maternal serum measurement at weeks 16 and 27		
Estradiol	0.72 [b]	0.61 [b]
Estriol	0.37 [b]	0.20 [a]
SHBG	0.85 [b]	0.67 [b]
Progesterone	0.45 [b]	0.39 [b]
Testosterone	0.83 [b]	0.62 [b]
Adiponectin	0.46 [b]	0.44 [b]
IGF-1	0.25 [b]	0.32 [b]
IGFBP3	0.16 [a]	0.32 [b]
Maternal serum measurement at weeks 16 with cord blood measurement		
Estradiol	0.09	0.15
Estriol	-0.04	-0.06
SHBG	0.09	-0.09
Progesterone	-0.001	0.03
Testosterone	-0.001	0.05
Adiponectin	0.18	-0.12
IGF-1	0.02	-0.02
IGFBP3	0.19	0.01
Maternal serum measurement at weeks 27 with cord blood measurement		
Estradiol	0.17	0.12
Estriol	0.22 [a]	0.00
SHBG	0.03	-0.12
Progesterone	0.09	0.06
Testosterone	0.16	0.17
Adiponectin	0.06	-0.20 [a]
IGF-1	0.18	0.29 [a]
IGFBP3	-0.02	-0.04

SHBG: sex hormone binding globulin; IGF-1: insulin-like growth factor 1; IGFBP3: insulin-like growth factor binding protein 3

[a] $p<0.05$, [b] $p<0.001$.

Table CP4.4. Comparison (%)[a] of the indicated hormones in maternal blood at the 16th and 27th gestational week and in the cord blood between Boston, USA (reference) and Shanghai, China

	16th gestational week[b]		**27th gestational week[b]**		**Cord blood[c]**	
	Shanghai vs. Boston	P	Shanghai vs. Boston	p	Shanghai vs. Boston	p
Estradiol (ng/ml)	23.0 (10.9, 36.5)	<0.001	7.2(-2.1, 17.3)	0.127	44.2 (12.5, 84.8)	0.004
Estriol (ng/ml)	34.9 (22.2, 49.0)	<0.001	42.9 (31.7, 55.0)	<0.001	2.6 (-13.2, 21.2)	0.775
SHBG (nmol/l)	4.2 (-1.7, 10.5)	0.170	-0.5 (-6.6, 6.0)	0.888	104.6 (47.2, 184.5)	<0.001
Prolactin (µg/l)	30.3 (14.9, 47.9)	<0.001	18.3 (8.1, 29.5)	<0.001	-	
Progesterone (ng/ml)	-3.2 (-8.7, 2.7)	0.274	-10.7 (-16.1, -4.8)	0.001	-20.4 (-40.1, 5.8)	0.120
Testosterone (ng/ml)	0.9 (-10.7, 14.0)	0.895	13.8 (1.3, 27.4)	0.023	54.5 (22.2, 95.4)	<0.001
Adiponectin (µg/ml)	-11.5(-20.2, -1.8)	0.022	-23.8 (-31.3, -15.5)	<0.001	-26.9 (-43.0, -6.3)	0.015
IGF-1 (ng/ml)	-36.9 (-45.4, -27.10)	<0.001	-14.9 (-24.7, -3.9)	0.011	-36.8 (-52.2, -16.3)	0.002
IGF-2 (ng/ml)	-		-		22.7 (9.2, 38.0)	0.001
IGFBP3 (ng/ml)	12.2 (2.52, 22.80)	0.013	0.5 (-9.0, 11.0)	0.937	33.7 (0.1, 78.4)	0.049

SHBG: sex hormone binding globulin; IGF: insulin-like growth factor; IGFBP3: insulin-like growth factor binding protein 3

[a] Hormone levels were log-transformed; the coefficient expresses % difference between centers.

[b] Controlling for maternal age, BMI, height and weight gain, parity, duration of gestation, exact gestational week and gender of offspring.

[c] Controlling as in [b] except for exact gestational week.

NA: Not applicable

We examined the possible concordance of the relative levels of these hormones with the sharp ecological contrast in the two populations with respect to breast cancer incidence. Results concerning maternal levels of estrogens, SHBG, prolactin and progesterone had been previously published (Brit J Cancer 1999; 79: 7-12), as had results concerning umbilical cord blood levels of IGF-1, IGF-2 and IGFBP-3 (Lagiou et al, 2009 – see Appendix 2). In contrast, hormone determinations conducted in the ILAT Steroid RIA Laboratory of the University of Massachusetts Medical School for maternal testosterone, adiponectin, IGF-1 and IGFBP-3, as well as cord blood estradiol, estriol, SHBG, progesterone, testosterone and adiponectin had not been previously reported.

We calculated three Spearman correlation coefficients for each endocrine compound, between the 16th week maternal sera measurement, the 27th week maternal sera measurement and the cord blood measurement. We also used multiple regression to compare log transformed cord blood and maternal serum hormone levels measured at the 16th and 27th gestational week between Boston (referent) and Shanghai, controlling for maternal age, maternal height, BMI before pregnancy and weight gain, as well as parity, duration of gestation, exact gestational week for maternal sampling and gender of offspring.

Spearman correlation coefficients between the levels of the hormones measured in maternal blood at the 16th and 27th gestational week and the levels of these hormones in the cord blood in the two settings are shown in Table CP4.3. Since correlation coefficients are dimensionless it is acceptable to calculate these coefficients, even when the measurements are done at different laboratories with different methods, provided the rankings within laboratories are not very different. It is evident that the measured hormones track during pregnancy, so that levels at the 16th gestational week are predictive of levels at the 27th gestational week, and vice versa. The correlations are strong with respect to SHBG, estradiol and testosterone, modest with respect to progesterone and adiponectin and weaker with respect to estriol, IGF-1 and IGFBP-3. In contrast, there was little or no correlation between hormone levels in maternal blood and levels of these hormones in cord blood.

In Table CP4.4, the levels of hormones measured in maternal blood at the 16th and 27th gestational week and in the cord blood are compared between Boston (reference) and Shanghai. Differences are evident between centers for most of the hormones in maternal blood or in the cord blood. In maternal sera, levels of estriol and prolactin (during both the 16th and 27th gestational week), as well as estradiol and IGFBP-3 (only during the 16th gestational week), and testosterone (only during the 27th gestational week) were significantly higher in Shanghai than in Boston, whereas the opposite was true for levels of IGF-1 and adiponectin (during both the 16th and 27th gestational week), as well as progesterone (only during the 27th gestational week); no significant differences were evident between the two centers with respect to SHBG at either gestational week. In the cord blood, levels of estradiol, testosterone, IGF-2, IGFBP-3 and most strikingly SHBG were significantly higher in Shanghai, whereas the opposite was true for levels of adiponectin and IGF-1.

We have attempted to interpret our findings in light of the considerably higher incidence of breast cancer in Caucasian compared to Chinese women, taking into account what is currently known about the role of these hormones in breast cancer etiology. We have considered differences in cord blood levels as more relevant to the fetal endocrine environment and the possible long term breast cancer risk of the offspring. In adult life estrogens and testosterone have been consistently positively associated with breast cancer risk. Both maternal and cord blood levels of estradiol, estriol and testosterone, however, were higher in Shanghai than in Boston by anywhere between 0.9% to 54.5% (Table CP4.4). Thus, levels of these pregnancy hormones by themselves are unlikely to be critical actors in the intrauterine origin of breast cancer since, if anything, they are higher in the population with substantially lower breast cancer incidence. In contrast, levels of SHBG were sharply higher in the cord blood of women in Shanghai (by 104.64%), raising the possibility that these high levels in Chinese women may reduce the bioavailability of active endogenous estrogens and testosterone in the fetus. In adult life, SHBG in relation to breast cancer has been studied mainly as a modulator of the effects of estradiol and testosterone (Mol Cell Endocrinol 2010; 316: 86-92), but there is also evidence that high levels of this compound may be more directly associated with a reduction in breast cancer risk (Epidemiology 1996; 7: 96-100).

In conclusion, taking into account the lower incidence of breast cancer among Chinese compared to Caucasian women and our current understanding of the role of the examined hormones in breast cancer risk, the endocrine factors likely to be involved in the intrauterine origin of breast cancer in the offspring are SHBG and IGF-2, with higher cord blood levels among Chinese, and IGF-1 with higher cord blood levels among Caucasian women.

CP4 reportable outcomes

- Results on the cord blood levels of IGF-1 and IGF-2 among Caucasian and Asian women and their association with birth size were presented in the February 2009, 2010 and 2011 LINKS meetings. Results on the cord blood levels of SHBG among Caucasian compared to Asian pregnancies were presented in the 2011 LINKS meeting.
- A manuscript on the cord blood levels of IGF-1 and IGF-2 among Caucasian and Asian women and their association with birth size, as a correlate of adult life breast cancer risk, has been published (Lagiou et al, 2009 – see Appendix 2)
- A manuscript on cord blood levels of SHBG and other hormones among Caucasian compared to Asian pregnancies, in the context of the sharp ecological contrast in breast cancer incidence in these populations, has been published (Lagiou et al, 2010 – see Appendix 2).
- A manuscript on the association between maternal energy intake and offspring gender has been published, with explicit acknowledgment of support by the Innovator project (see Appendix 4)

CP4 Conclusion

The work under this component project was successfully performed. Two manuscripts have been published and further analyses are considered.

CP4 References

Lagiou P, Hsieh CC, Lipworth L, Samoli E, Okulicz W, Troisi R, Xu B, Hall P, Ekbom A, Adami HO, Trichopoulos D. Insulin-like growth factor levels in cord blood, birth weight and breast cancer risk. Br J Cancer. 2009;100:1794-8.

Lagiou P, Samoli E, Okulicz W, Xu B, Lagiou A, Lipworth L, Georgila C, Vatten L, Adami HO, Trichopoulos D, Hsieh CC. Maternal and cord blood hormone levels in the United States and China and the intrauterine origin of breast cancer. Ann Oncol. 2010 Oct 13. [Epub ahead of print]

CP5: "Breast stem cells and perinatal factors for breast cancer risk"

CP5 PI: Prof. Chung-Cheng Hsieh, University of Massachusetts Cancer Center, 55 Lake Avenue North, Worcester, MA 01655

Timetable of research accomplishments of CP5 as outlined in the Statement of Work.

Task 5 To investigate whether markers of mammary stem cells are associated with perinatal characteristics that are linked to breast cancer in later life:

a. Finalization of questionnaire for obtaining maternal and gestation characteristics. (Months 1-3)
b. Training of the study personnel on study procedures. (Months 1-6)
c. Subject recruitment and sample collection from a total of 250 pregnant women. (Months 7-42)
d. Conduct of laboratory assays for markers of stem cells (Months 7-45)
e. Conduct of laboratory assays for hormones (Months 10-48)
f. Data analyses. (Months 45-54)
g. Manuscript preparation and submission. (Months 49-60)

CP5 Progress report

The underlying premises of this component project are that 1) breast stem cells are the cell type that undergoes malignant transformation, 2) breast stem cells primarily arise during the fetal/perinatal period, and therefore the *in utero*/perinatal environment is a major determinant of the breast stem cell number in an individual, and 3) the greater the number of breast stem cells,

the greater the likelihood that one will undergo an oncogenic alteration. Previously, we have shown that the concentration of hematopoietic stem cells in umbilical cord blood, serving as a surrogate for overall stem cell levels, is correlated with perinatal levels of mitogens including estrogens and IGF-1 (Cancer Res 2005; 65; 358-363; Breast Cancer Res 2007; 9: R29). Ideally, one would like to obtain some indicators of the levels of epithelial precursors cells in the perinatal environment, and determine if such cells might be a better indicator of future breast cancer risk. In CP5, we are using umbilical cord blood as the perinatal cell source.

Ethical clearance of CP5 (a+b) was granted by the US Army Human Subjects Research Review Board (HSRRB) on July 17, 2006 and CP5 was officially launched on July 18, 2006. The questionnaire for obtaining maternal and gestation characteristics has been finalized and study personnel have been trained on the study procedures (Task 5a and 5b).

Subject recruitment and sample collection started at the Tufts-New England Medical Center (currently known as Tufts Medical Center) on November 16, 2006 (Task 5c). From November 16, 2006 to the termination of recruitment for this study on November 17, 2010, 287 women consented to participate in the study. Umbilical cord blood samples were collected from 216 eligible subjects. Of these samples, 179 were processed at the University of Massachusetts Cancer Center Stem Cell Research Laboratory within 24 hours of childbirth (Task 5d). After the unexpected passing away of Dr. Todd Savarese in October 2007, a laboratory scientist, Dr. Li Qiu, was recruited in May 2008 to work on the project.

In the past grant year (no cost extension), the research team at the University of Massachusetts Cancer Center has continued to make progress by processing 21 umbilical cord blood samples of normal pregnancies received from the Tufts Medical Center for mononuclear cells (MNC) (Task 5d). Genes of putative epithelial breast stem/progenitor cells, such as epithelial adhesion molecule (EpCAM), CD49f (α6-integrin), CD117 (c-kit receptor), CD24, and CD29 (β1-integrin), were detected in umbilical cord blood-derived MNC by reverse transcription polymerase chain reaction (Figure CP5.1A). Additionally, we confirmed co-localized protein expressions of $EpCAM^+CD49f^+$, $CD49f^+CD24^+$ and $CD24^+CD29^+$ surface markers in umbilical cord blood-derived MNC by confocal microscopic analyses. We further quantified the percentages of MNC derived from umbilical cord blood by flow cytometry for hematopoietic cell markers ($CD34^+$ and $CD34^+CD38^-$) and reported markers of breast stem/progenitor cells ($EpCAM^+CD49f^+$, $EpCAM^+CD49f^+CD117^+$, $CD49f^+CD24^+$, $CD24^+CD29^+$, and $CD24^+CD29^+CD49f^+$), in addition to the general epithelial marker EpCAM. Data analyses (Task 5f) using the FlowJo software program showed that the $EpCAM^+$ subpopulation ranged from 0.19 to 19.8 cells/1,000 MNC with a mean±standard deviation of 3.4 ± 4.0 cells; the $EpCAM^+CD49f^+$ subpopulation ranged from 0.049 to 9.7 cells/1,000 MNC with a mean of 1.7 ± 1.8 cells; the $EpCAM^+CD49f^+CD117^+$ subpopulation ranged from 0.02 to 2.4 cells/1,000 MNC with a mean of 0.48 ± 0.56 cells; the $CD49f^+CD24^+$ subpopulation ranged from 0 to 48.1 cells/1,000 MNC with a mean of 14.7 ± 12.9 cells; the $CD24^+CD29^+$ subpopulation ranged from 0.11 to 46.2 cells/1,000 MNC with a mean of 10.3 ± 9.8 cells; and the $CD24^+CD29^+CD49f^+$ subpopulation ranged from 0 to 44.4 cells/1,000 MNC with a mean of 8.3 ± 8.8 cells. We performed a rank correlation analysis between concentrations of hematopoietic and breast stem/progenitor cell subpopulations and found that levels of the $EpCAM^+$ subpopulation were

positively correlated with concentrations of $CD34^+$ and $CD34^+CD38^-$ hematopoietic stem cells (both $P = 0.006$). Except for the $CD24^+CD29^+$ cells, all putative breast stem/progenitor cell subpopulations were positively associated with the hematopoietic stem cell subpopulations. Notably, the $EpCAM^+CD49f^+$ and $CD49f^+CD24^+$ subpopulations were positively and significantly correlated to the $CD34^+$ cells ($P = 0.03$ and 0.008, respectively). These associations were clearer among female than among male newborns, in particular the $EpCAM^+$ subpopulation. A manuscript reporting these findings has been accepted for publication in Annals of Oncology (Task 5g). Additionally, plasma harvested from these samples has been sent for hormone assays (Task 5e).

We have not been able to establish parameters for the liquid culture of putative breast stem cell-derived cell clusters ("mammospheres"). While a variety of culture conditions [including the growth of MNC or $EpCAM^+$ MNC in defined mammary epithelial cell growth medium (MEGM) on Collagen-1 coated dishes; the growth of MNC or lineage-depleted MNC in defined mammocult medium on tissue culture or ultra-low attachment dishes; and the growth of MNC in mammocult medium supplemented with fetal bovine serum, β-mercaptoethanol, or sonic hedgehog] resulted in the formation of cell clusters, inconsistent growth and cell death within 1 to 2 weeks made it difficult to assess clonal growth. Since the last report, we have investigated the growth of differentiated mammary phenotypes by culturing MNC, $EpCAM^+$ and $CD49f^+$ MNC in matrigel but with no success. Recently we have some success in obtaining putative mammary phenotypes by culturing $CD10^+$ MNC in matrigel that needs to be confirmed by immuno-staining.

CP5 key research accomplishments

Summary:

- Among term newborns with a normal-to-high birth weight, birth weight was found to be significantly positively associated with stem cell measurements, supporting a role of stem cell pool on cancer risk (Strohsnitter et al, 2008 – see Appendix 2).
- Genes of putative epithelial breast stem/progenitor cell markers, i.e., EpCAM, CD49f (α6-integrin), CD117 (c-kit), CD24 and CD29 (β1-integrin), were detected in umbilical cord blood-derived MNC by RT-PCR (Qiu et al, 2011 – see Appendix 2).
- Protein expressions of putative epithelial breast stem/progenitor cells, i.e., $EpCAM^+CD49f^+$, $CD49f^+CD24^+$ and $CD24^+CD29^+$ surface markers, were detected in umbilical cord blood-derived MNC by immuno-staining and confocal microscopic analyses (Qiu et al, 2011– see Appendix 2).
- Subpopulations with markers for hematopoietic stem cells ($CD34^+$ and $CD34^+CD38^-$) and putative epithelial breast stem/progenitor cells ($EpCAM^+$, $EpCAM^+CD49f^+$, $EpCAM^+CD49f^+CD117^+$, $CD49f^+CD24^+$, $CD24^+CD29^+$, and $CD49f^+CD24^+CD29^+$) were detectable in umbilical cord blood-derived MNC and quantified by flow cytometric analysis (Qiu et al, 2011– see Appendix 2).
- Umbilical cord blood breast and hematopoietic stem cells have been found to be positively correlated ($EpCAM^+$ with $CD34^+$ and $CD34^+CD38^-$ cells; $EpCAM^+CD49f^+$ and $CD49f^+CD24^+$ with $CD34^+$ cells) (Qiu et al, 2011– see Appendix 2).

Specifically:

In a paper published in the British Journal of Cancer in 2008 in the context of CP5 (Strohsnitter et al, 2008 – see Appendix 2), we examined the association of the size of the hematopoietic stem cell pool with birth weight. The paper was based on the established positive association of birth weight with the risk of breast cancer in adult life (PLoS Med 2008; 5: e193) and the hypothesis (supported in a later study in the context of CP5 – see below) that the size of the hematopoietic stem cell pool is correlated to the size of the mammary tissue specific stem cell pool. The postulated underlying pathophysiologic mechanism was that levels of in utero/perinatal mitogens and other factors determine the size of the stem cell pools in the developing fetus, and the greater the stem cell pool size, the greater the chance that one of the stem cells will be mutated by a carcinogen, or undergo a DNA replicative error, initiating oncogenic transformation.

Consenting study subjects were recruited from one of two sources: (1) participants in the Worcester, MA-based American Red Cross cord blood program (ACBP), in which haematopoietic stem cells from umbilical cord blood were collected for possible transplantation, from August 2002 to June 2003, and (2) pregnant women delivering at T-NEMC from October 2004 to April 2006. All the cord blood samples were from full-term singleton infants. The processing of samples included the determination of cord blood volumes, the determination of initial levels of total nucleated cells (TNC) and mononuclear cells (MNC) before centrifugations or manipulation, the quantitation of haematopoietic stem / progenitor cell populations and the determination of cord blood plasma hormone levels (Breast Cancer Res 2007;9: R29). The following haematopoietic stem and progenitor populations were quantitated (1) $CD34^+$ cells, a heterogeneous population of early multipotent stem and progenitor cells, committed progenitors and differentiating cells (Leuk Lymphoma 2000;38:489–497); (2) $CD34^+CD38^-$ cells, which represent more primitive stem cells depleted of lineage-committed precursors (Leuk Lymphoma 2000;38:489–497); (3) $CD34^+$c-*kit*$^-$ cells, which also represent a more primitive stem cell population that has relatively high cloning efficiencies in semisolid culture (Cytokine 1994;6: 195–205 ; Stem Cells 1998;16: 153–165); and (4) granulocyte–macrophage colony forming units (CFU-GM), a functional measure of the number of proliferative granulocyte / macrophage committed haematopoietic precursor cells (Exp Hematol 1992;20:1043–1047).

Multivariate linear regression was used to examine the association between natural log-transformed measures of stem cell potential (dependent variable) and birth weight (independent variable, using 3000–3499 g as the reference), adjusting for maternal and neonatal characteristics (mother's age, race of parents, number of live births, gestation duration, baby's gender, delivery time and study site). To assess whether there was an underlying linear trend, birth weight was next analyzed as a continuous variable across the whole range of birth weight values with the effect estimates expressed for each 500 g increase in birth weight. The fitted coefficients from the regression analyses were exponentiated to obtain the estimated proportional change in birth weight associated with each independent variable. Statistical significance was set at 0.05 (two-sided).

We found a J-shaped association between birth weight categories and concentrations of TNC (lymphocytes, monocytes and granulocytes), as well as a J-shaped relation with MNC (lymphocytes and monocytes), both including more differentiated cells. Among the stem cell populations, there was a positive association with $CD34^{+}CD38^{-}$ cells across the whole range of birth weight categories, with each 500 g increase being associated with 15.5% higher levels of this cell population (95% confidence interval: 1.6, 31.3%) (Table CP5.1).

A J-shaped relation was observed for the $CD34^{+}$ and $CD34^{+}$c-*kit*$^{-}$ cells: for birth weights of 3000 g or greater, stem cell concentrations increased with birth weight, while the lowest category of o3000 g had higher levels than the category of 3000–3499 g. For CFU-GM, an approximate U-shaped relation was observed, with the lowest birth weight category having the highest levels of this cell population (Table CP5.1).

Table CP5.1. Multiple linear regression analysis for the association between measurements of hematopoietic stem cell populations and birth weight.

Analytic model	Cell measurements	Birth weight (g) in categories				Birth weight per 500g
		<3,000	3,000-3,499	3,500-3,999	>=4,000	
		% difference (95% CI)	Reference (geometric mean*)	% difference (95% CI)	% difference (95% CI)	% difference (95% CI)
Adjusted for core covariates**	TNC	2.1 (-8.5, 14.0)	0.0 (15.00)	7.8 (-1.7, 18.3)	8.4 (-8.1, 27.8)	3.0 (-2.5, 8.9)
	MNC	5.7 (-6.4, 19.3)	0.0 (6.91)	6.5 (-3.9, 18.0)	12.5 (-6.2, 34.9)	2.1 (-3.9, 8.4)
	$CD34^{+}$ [a]	2.4 (-19.2, 29.7)	0.0 (7.04)	12.5 (-8.1, 37.7)	20.8 (-15.5, 72.8)	9.5 (-2.6, 23.2)
	$CD34^{+}CD38^{-}$ [a]	-1.1 (-23.6, 28.1)	0.0 (3.11)	19.1 (-4.3, 48.1)	**47.9 (0.4, 117.8)**	**15.5 (1.6, 31.3)**
	$CD34^{+}$c-*kit*$^{+}$ [b]	13.7 (-14.0, 50.4)	0.0 (5.80)	18.9 (-5.5, 49.6)	34.2 (-10.4, 101.1)	10.0 (-4.0, 26.1)
	CFU-GM[c]	**37.0 (2.1, 83.8)**	0.0 (4.04)	29.0 (-0.1, 66.6)	31.4 (-16.2, 106.1)	6.1 (-8.4, 23.0)

*Unadjusted geometric means. TNC: initial total nucleated cells x 10^6/ml; MNC: initial mononuclear cells x 10^6/ml; the unit for the stem cell populations ($CD34^{+}$, $CD34^{+}CD38^{-}$, $CD34^{+}$c-*kit*$^{+}$, and CFU-GM) was /1,000 MNC.

**Core covariates included mother's age, race of parents (both Caucasian or not), parity, gestation duration, gender of baby (male or female), delivery time (night or day), and study site (ACBP or T-NEMC).

[a] n=233 with complete information on all the covariates

[b] Determined only in the T-NEMC-derived samples

[c] Data from the T-NEMC-derived samples on which this assay was conducted

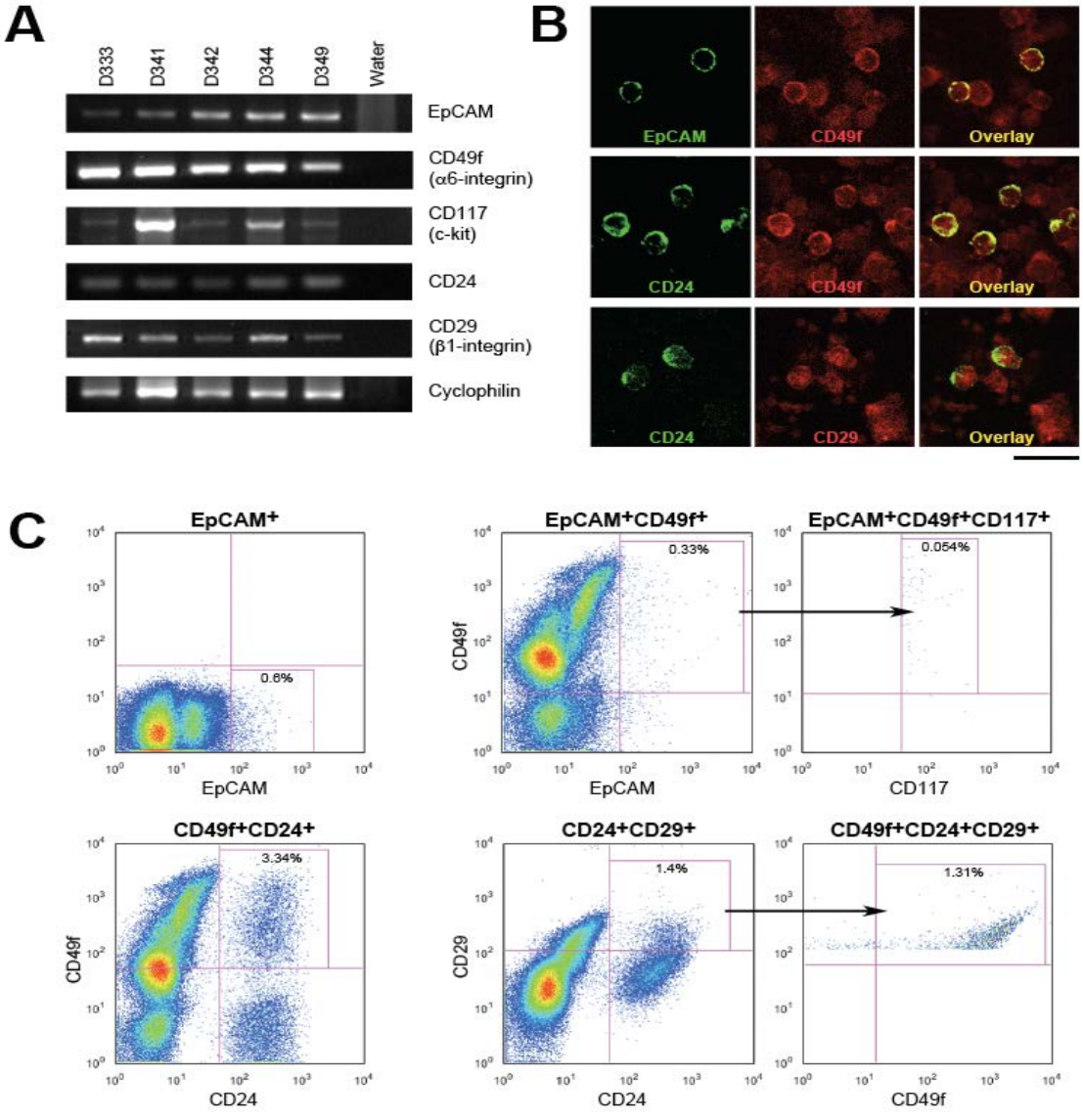
Our findings indicate that there is a positive association between birth weight and haematopoietic stem cell measurements in the cord blood samples among newborns with normal-to-high (3000 g or more) birth weights. This association is strongest with $CD34^{+}CD38^{-}$ cells, a relatively primitive haematopoietic stem cell population. These data are in line with previous studies, which showed a positive relationship between cord blood $CD34^{+}$ or CFU-GM levels and birth weight (Br J Haematol 1998;103: 1167–1171 ; Bone Marrow Transplant 2001;27:7–14 ; Acta Paediatr 2004;93:1323–1329). We conclude that there is a J- or U-shaped positive association between birth weight and stem cell measurements, linear among term newborns with normal-to-high birth weight and stronger with the more primitive stem cell sub-population. The nonlinear relationship, however, suggests that birth weight is not an unequivocal indicator of stem cell potential in the context of a prenatal origin of cancer risk. Quantitation of stem cell pools might prove to be a more accurate predictor of cancer risk than birth weight per se.

In another paper in the context of CP5, accepted for publication by the Annals of Oncology (Qiu et al, *in press* – see Appendix 2), we analyzed mononuclear cells derived from human umbilical cord blood for gene expressions of markers reported for putative breast epithelial stem cells and progenitors (Genes Dev 2009; 23: 2563-2577) and quantified such cell populations by flow cytometry. We were able to document the existence of putative breast stem/progenitor cell phenotypes in umbilical cord blood and that the concentrations of certain breast stem/progenitor cell phenotypes found in umbilical cord blood correlate positively with those of hematopoietic stem cells.

Study subjects were recruited from November, 2006 to November, 2010 among pregnant women who delivered at the Tufts Medical Center, who were 18 years or older, with a fetus free of anomalies by ultrasound examination. We included only full term, normotensive, singleton pregnancies.

Umbilical cord blood was collected from the umbilical vein into a sterile bag containing 35 ml of citrate-phosphate-dextrose anticoagulant (Fenwal, Lake Zurich, IL, USA). Samples were processed for umbilical cord blood mononuclear cells (MNC) using a Ficoll-Paque density gradient within 24 hours of birth. We determined, by reverse transcription polymerase chain reaction, the presence of genes of putative breast epithelial stem cell/progenitor markers (including EpCAM, CD49f (α6-integrin), CD117 (c-kit), CD24, and CD29 (β1-integrin)). By immunocytochemistry, we co-localized protein expressions of $EpCAM^{+}CD49f^{+}$, $CD49f^{+}CD24^{+}$ and $CD24^{+}CD29^{+}$. Using Spearman rank correlation coefficients, we correlated concentrations of putative breast stem cell/progenitor subpopulations, quantified by flow cytometry, with concentrations of hematopoietic stem cells.

Figure CP5.1. (A) Gel electrophoresis showing the detection of PCR products for EpCAM, CD49f (α6-integrin), CD117 (c-kit), CD24, CD29 (β1-integrin), and the housekeeping gene cyclophilin from umbilical cord blood-derived mononuclear cells from 5 umbilical cord blood samples (D333, D341, D342, D344, and D349). Water (last lane) was used as negative controls. (B) Double-labeled immunofluorescent confocal microscopy of umbilical cord blood-derived mononuclear cells from sample D321 showing colocalization in the overlay image of EpCAM (green) and CD49f (red) (top panel); CD24 (green) and CD49f (red) (middle panel); and CD24 (green) and CD29 (red) (bottom panel). Scale bar represents 20 μm. (C) Flow cytometric pseudocolor plots showing the detection of the $EpCAM^+$, $EpCAM^+CD49f^+$, $EpCAM^+CD49f^+CD117^+$, $CD49f^+CD24^+$, $CD24^+CD29^+$, and $CD49f^+CD24^+CD29^+$ subpopulations (boxed, with percentage of cells indicated) from umbilical cord blood-derived mononuclear cells of sample N41. The arrows indicate that the triple positive population was derived from the double positive population as shown. The markers shown in the top panel have been reported in humans while the markers shown in the bottom panel have been reported in mice.



More specifically, to determine whether putative breast stem/progenitor cell phenotypes were present in umbilical cord blood, we analyzed genes reported for breast stem cell markers (Genes Dev 2009; 23: 2563-2577) in the MNC fraction of umbilical cord blood by RT-PCR. Because breast stem cells are considered epithelial in nature (Stem Cell Rev 2006; 2: 103-110), we first detected the gene for epithelial adhesion molecule (EpCAM), or epithelial-specific antigen (ESA), as a marker for epithelial cells (J Cell Biol 1994; 125: 437-446). Additionally, genes of putative markers for breast stem/progenitor cells, i.e., CD49f (α6-integrin), CD117 (c-kit receptor), CD24, and CD29 (β1-integrin) were detected in umbilical cord blood-derived MNC (Figure CP5.A).

Second, we determined protein expressions by immunocytochemistry and observed co-localized staining of $EpCAM^+CD49f^+$, $CD49f^+CD24^+$ and $CD24^+CD29^+$ surface markers in umbilical cord blood-derived MNC by confocal microscopic analyses (Figure CP5.B). We further quantified the percentages of umbilical cord blood-derived MNC with putative markers of the different breast stem cell subpopulations ($EpCAM^+CD49f^+$, $EpCAM^+CD49f^+CD117^+$, $CD49f^+CD24^+$, $CD24^+CD29^+$, and $CD24^+CD29^+CD49f^+$) in addition to EpCAM by flow cytometry (Figure CP5.C).

Data analyses using the FlowJo software program showed that the $EpCAM^+$ subpopulation ranged from 0.19 to 19.8 cells/1,000 MNC with a mean ± standard deviation of 3.4 ± 4.0 cells; the $EpCAM^+CD49f^+$ subpopulation ranged from 0.049 to 9.7 cells/1,000 MNC with a mean of 1.7 ± 1.8 cells; the $EpCAM^+CD49f^+CD117^+$ subpopulation ranged from 0.02 to 2.4 cells/1,000 MNC with a mean of 0.48 ± 0.56 cells; the $CD49f^+CD24^+$ subpopulation ranged from 0 to 48.1 cells/1,000 MNC with a mean of 14.7 ± 12.9 cells; the $CD24^+CD29^+$ subpopulation ranged from 0.11 to 46.2 cells/1,000 MNC with a mean of 10.3 ± 9.8 cells; and the $CD24^+CD29^+CD49f^+$ subpopulation ranged from 0 to 44.4 cells/1,000 MNC with a mean of 8.3 ± 8.8 cells.

We also quantified the percentages of umbilical cord blood-derived MNC with hematopoietic stem cell markers, i.e., $CD34^+$ and $CD34^+CD38^-$, and performed a rank correlation analysis between concentrations of hematopoietic and breast stem/progenitor cell subpopulations. Levels of the $EpCAM^+$ subpopulation were positively correlated with concentrations of $CD34^+$ and $CD34^+CD38^-$ hematopoietic stem cells (both $P = 0.006$; Table CP5.2). Except for the $CD24^+CD29^+$ cells in female offspring, all putative breast stem/progenitor cell subpopulations were positively associated with the hematopoietic stem cell subpopulations. Notably, the $EpCAM^+CD49f^+$ and $CD49f^+CD24^+$ subpopulations were positively and significantly correlated to the $CD34^+$ cells ($P = 0.03$ and 0.008, respectively). These associations were clearer among female than among male newborns, in particular for the $EpCAM^+$ subpopulation (Table CP5.2).

Table CP5.2. Spearman correlation coefficients (*P* values in parentheses) between umbilical cord blood hematopoietic and breast stem/progenitor cell populations.

Hematopoietic stem cell subpopulations	Breast stem/progenitor cell subpopulations					
	$EpCAM^+$	$EpCAM^+CD49f^+$	$EpCAM^+CD49f^+CD117^+$	$CD49f^+CD24^+$	$CD24^+CD29^+$	$CD49f^+CD24^+CD29^+$
All subjects (*N*)	112	109	32	56	55	55
$CD34^+$	0.26 (**0.006**)	0.21 (**0.03**)	0.24 (0.20)	0.35 (**0.008**)	0.14 (0.32)	0.17 (0.21)
$CD34^+CD38^-$	0.26 (**0.006**)	0.15 (0.12)	0.30 (0.09)	0.24 (0.07)	0.01 (0.97)	0.18 (0.19)
Newborn gender						
Males (*N*)	53	53	14	25	25	25
$CD34^+$	0.23 (0.37)	0.22 (0.12)	0.41 (0.15)	0.33 (0.11)	0.07(0.76)	0.28 (0.18)
$CD34^+CD38^-$	0.02 (0.90)	0.04 (0.77)	0.27 (0.36)	0.38 (0.06)	0.09 (0.67)	0.36 (0.08)
Females (*N*)	59	56	18	31	30	30
$CD34^+$	0.36 (**0.005**)	0.18 (0.19)	0.03 (0.90)	0.35 (**0.05**)	0.18 (0.34)	0.05 (0.78)
$CD34^+CD38^-$	0.37 (**0.004**)	0.19 (0.17)	0.18 (0.48)	0.19 (0.29)	-0.03 (0.88)	0.08 (0.68)

To our knowledge, this is the first report of measurable breast stem cells in human umbilical cord blood. The detection of putative breast stem/progenitor cell phenotypes suggests that there is a "mammary" compartment within the umbilical cord blood. These putative breast stem/progenitor cell phenotypes in the umbilical cord blood niche are rare, but detectable by flow cytometry. Our results indicate that, at the time of birth, intrauterine conditions may sustain certain subpopulations of breast stem/progenitor cell phenotypes.

Levels of these breast stem cells correlated positively and significantly with those of hematopoietic stem cells. In previous studies we had shown that concentrations of hematopoietic stem cells (as surrogate measurements of overall stem cell levels in the intrauterine environment) are correlated with umbilical cord blood levels of hormones and with birth weight (Cancer Res 2005; 65:358-363 ; Breast Cancer Res 2007;9: R29 ; Br J Cancer 2008; 98: 660-663). Thus, our results provide a biologically plausible mechanism underlying the association of early life exposures with the risk of breast cancer in the offspring through a pathway involving the mammary stem cell pool.

CP5 reportable outcomes

- Recruitment of 287 pregnant women has been accomplished.
- 179 umbilical cord blood samples have been processed for stem cell measurements.
- The CP5 co-PI (CCH) attended the Era of Hope 2008 meeting and presented results on the association of birth weight with stem cell measurements.
- A paper on the positive association of hematopoietic stem cell measurements with birth weight has been published (Strohsnitter et al, 2008 – see Appendix 2).
- A paper on the identification of mammary stem cell phenotypes in umbilical cord blood and the positive correlation of hematopoietic stem cells with mammary tissue specific stem cells has been accepted for publication by the Annals of Oncology journal (Qiu et al, *in press* – see Appendix 2)
- Results on the positive association of birth weight with hematopoietic stem cell measurements were presented in the February 2010 LINKS meeting and results on the positive correlation of hematopoietic stem cells with mammary tissue specific stem cells were presented in the 2011 LINKS meeting.

CP5 Conclusion

This is the first report on the identification of measurable mammary stem cell phenotypes in human umbilical cord blood. Importantly, the concentrations of certain breast stem/progenitor cell populations found in umbilical cord blood correlated positively and significantly with those of hematopoietic stem cells. This finding is significant because concentrations of hematopoietic stem cells have been correlated with indicators of adult life breast cancer risk, notably birth weight, as well as umbilical cord blood plasma levels of IGF-1 and estrogens. Hence, breast stem cell pools in umbilical cord blood are likely to be predictors of breast cancer susceptibility in the adult life. However, future research will need to determine whether these putative breast stem/progenitor cell pools identified in umbilical cord blood are indeed functional mammary

cells. If so, we will have a model system to understand whether genetic and/or epigenetic alterations in breast stem/progenitor cells explains fetal programming of breast cancer risk in the adult life.

CP5 References

Strohsnitter WC, Savarese TM, Low HP, Chelmow DP, Lagiou P, Lambe M, Edmiston K, Liu Q, Baik I, Noller KL. Adami HO, Trichopoulos D, Hsieh CC. Correlation of umbilical cord blood haematopoietic stem and progenitor cell levels with birth weight: implications for a prenatal influence on cancer risk. Br J Cancer. 2008;98:660-3.

Qiu L, Low HP, Chang CI, Strohsnitter WC, Anderson M, Edmiston K, Adami HO, Ekbom A, Hall P, Lagiou P, Trichopoulos D, Hsieh CC. Novel measurements of mammary stem cells in human umbilical cord blood as prospective predictors of breast cancer susceptibility in later life. Ann Oncol. 2011; in press.

Task 6: "Monitoring, coordination and fine-tuning of the five component projects"

Timetable of accomplishments of task 6 as outlined in the Statement of Work.

Task 6 To monitor, coordinate and fine-tune the five component projects:

a. Continuous monitoring and coordination of the research activities under the five component projects. (Months 1-60)
b. Compilation of annual overall project reports based on the component project-specific annual reports. (annually throughout the duration of the project) (Months1-60)
c. To coordinate the preparation and submission of manuscripts produced (Months 37-60)

Task 6 progress report

Monitoring and coordination of the five component projects has presented no problems throughout the duration of the project (including the one year no cost extension granted). The key investigators have a long history of successful scientific collaboration, which continued in the context of the current project.

The project PI (DT) and the project coordinator (PL) attended the 2006 LINKS meeting and presented the outline of the W81XWH-05-1-0314 Innovator project.

In the context of his receiving the 2007 Medal of Honour by the International Agency for Research on Cancer, WHO, for his research on cancer, the project PI (DT) was invited to write a perspective on early life events and conditions and breast cancer risk. The perspective was co-authored by the PI and four Innovator key researchers. Support by the W81XWH-05-1-0314 Innovator Award was acknowledged. (Trichopoulos et al, 2008 – see Appendix 3)

The project coordinator (PL) and the project PI (DT) were invited to write an editorial on a metaanalysis confirming the association of birth size with breast cancer risk in which acknowledgement to the W81XWH-05-1-0314 Innovator award and reference to the role of early life influences on breast cancer risk was made (Lagiou and Trichopoulos, 2008 – see Appendix 3).

The component project 5 co-PI (CCH) attended the Era of Hope 2008 meeting and presented results from component project 5.

The project PI (DT) attended the 2009 LINKS meeting and presented results from component projects 1, 2 and 4.

The project PI (DT) attended the 2010 and 2011 LINKS meeting and presented results from all 5 component projects.

Four papers (Cancer Epidemiol Biomarkers Prev. 2007;16:1720-3 ; Cancer Epidemiol Biomarkers Prev. 2008;17:224-31 ; Breast Cancer Res. 2010 Mar 9;12(2):R19 ; Eur J Epidemiol. 2011;26:39-44) addressing issues indirectly relevant to the W81XWH-05-1-0314 Innovator Award have been published with acknowledgement to support by this Award (see Appendix 4).

Task 6 References

Lagiou P, Trichopoulos D. Birth size and the pathogenesis of breast cancer. PLoS Med. 2008;5(9):e194.

Trichopoulos D, Adami HO, Ekbom A, Hsieh CC, Lagiou P. Early life events and conditions and breast cancer risk: From epidemiology to etiology. Int J Cancer. 2008;122:481-485.

KEY RESEARCH ACCOMPLISHMENTS

- Postnatal weight loss (an indicator of water loss, likely to reflect water retention associated with pregnancy hormones), as well as neonatal weight gain rate after reaching the nadir (known to reflect growth hormone levels), are significantly positively associated with premenopausal breast cancer risk. (CP1)

- Birth size is associated with postmenopausal mammographic density measured more than five decades later. These results support the hypothesis that adult breast density, a powerful correlate of breast cancer risk, has intrauterine roots, as reflected in birth size. (CP2)

- There is suggestive evidence for interaction of birth weight with genetic predisposition in relation to breast cancer risk, particularly with respect to SNP rs2981582. (CP3)

- Cord blood levels of IGF-1 are substantially higher among Caucasian babies born in Boston compared to Chinese in Shanghai. The documentation of higher cord blood levels of IGF-1 - a principal growth hormone which does not cross the placenta - among Caucasian compared to Asian newborns is concordant with breast cancer incidence in these populations. (CP4)

- Cord blood levels of IGF-2 are substantially higher among Chinese babies born in Shanghai compared to Caucasian babies born in Boston. (CP4)

- IGF-1 is significantly positively associated with birth weight and birth length in Boston, but not Shanghai. Positive, though statistically non-significant, associations of IGF-2 with birth size are only evident in Shanghai. (CP4)

- For cord blood IGF-1 in Boston and for cord blood IGF-2 in Shanghai, the associations with birth weight and birth length are positive and significant among taller women (among whom maternal anthropometry is likely to impose fewer constrains) but essentially null among shorter women. (CP4)

- Levels of SHBG are sharply higher in the cord blood of women in Shanghai (by 104.64%), raising the possibility that these high levels in Chinese women may reduce the bioavailability of active endogenous estrogens and testosterone in the fetus. (CP4)

- Among term newborns with a normal-to-high birth weight, birth weight was found to be significantly positively associated with hematopoietic stem cell measurements, supporting a role of stem cell pool on cancer risk. (CP5)

- Genes of putative breast stem cell markers, such as EpCAM, CD49f (α-6 integrin), CD24 and CD29 (β-1 integrin) have been detected in umbilical cord blood-derived mononuclear cells (MNC) by RT-PCR. (CP5)

- Protein expressions of $EpCAM^{+}CD49f^{+}$, $CD49f^{+}CD24^{+}$ and $CD24^{+}CD29^{+}$ can be detected by immuno-staining and co-localizations are demonstrated by confocal microscopy of umbilical cord blood-derived MNC. (CP5)

- Subpopulations with markers for hematopoietic stem cells ($CD34^{+}$ and $CD34^{+}CD38^{-}$) and putative breast stem cells ($EpCAM^{+}CD49f^{+}$, $EpCAM^{+}CD49f^{+}CD117^{+}$, $CD49f^{+}CD24^{+}$, $CD24^{+}CD29^{+}$, and $CD24^{+}CD29^{+}CD49f^{+}$) can be detected by flow cytometric analysis of umbilical cord blood-derived MNC. To our knowledge, this is the first report of measurable breast stem cells in human umbilical cord blood. (CP5)

- Levels of these breast stem cells correlate positively and significantly with those of hematopoietic stem cells. (CP5)

REPORTABLE OUTCOMES

- A manuscript on the association of postnatal growth with breast cancer risk has been published (Lagiou et al, 2008 – see Appendix 2). (CP1)

- A manuscript on the association of birth size with postmenopausal mammographic density has been published (Tamimi et al, 2010a – see Appendix 2). (CP2)

- A manuscript on the interaction of birth weight with higher breast cancer risk SNPs in relation to breast cancer has been published (Tamimi et al, 2010b – see Appendix 2). (CP3)

- A manuscript on the cord blood levels of IGF-1 and IGF-2 among Caucasian and Asian women and their association with birth size has been published (Lagiou et al, 2009 – see Appendix 2) (CP4)

- A manuscript on cord blood levels of SHBG and other hormones among Caucasian compared to Asian pregnancies, in the context of the sharp ecological contrast in breast cancer incidence in these populations, has been published (Lagiou et al, 2010 – see Appendix 2).

- A manuscript on the association of stem cell measurements with birth weight has been published (Strohsnitter et al, 2008 – see Appendix 2). (CP5)

- A manuscript on the identification of mammary stem cell phenotypes in umbilical cord blood and the positive correlation of hematopoietic stem cells with mammary tissue specific stem cells has been accepted for publication by the Annals of Oncology journal (Qiu et al, *in press* – see Appendix 2)

- The project PI (DT) and the project coordinator (PL) attended the 2006 LINKS meeting and presented the outline of the W81XWH-05-1-0314 Innovator project. (CP6)

- The CP5 5 co-PI (CCH) attended the Era of Hope 2008 meeting and presented results on the association of birth weight with stem cell measurements. (CP6)

- The project PI (DT) attended the 2009 LINKS meeting and presented results from component projects 1, 2 and 4. (CP6)

- The project PI (DT) attended the 2010 and 2011 LINKS meeting and presented results from all 5 component projects. (CP6)

- The project PI (DT) was invited to write a perspective on early life events and conditions and breast cancer risk. The perspective was co-authored by the PI and four Innovator key researchers. Support by the W81XWH-05-1-0314 Innovator Award was acknowledged. (CP6 - see Appendix 3)

- The project coordinator (PL) and the project PI (DT) were invited to write an editorial on a metaanalysis confirming the association of birth size with breast cancer risk in which acknowledgement to the W81XWH-05-1-0314 Innovator award and reference to the role of early life influences on breast cancer risk was made. (CP6- see Appendix 3)

- Four manuscripts addressing issues indirectly relevant to the W81XWH-05-1-0314 Innovator Award have been published with acknowledgement to support by this Award. (see Appendix 4)

CONCLUSIONS

The hypothesis that breast cancer has intrauterine roots was elaborated in 1990 (Lancet 1990;335:939-940). Soon afterwards, it was argued that birth size and pregnancy toxemia may allow empirical testing of the hypothesis (Lancet 1992; 340:1015-1018). Independently of this Innovator project, but relying on this hypothesis, a major metaanalysis documented that higher birth size is associated with higher breast cancer risk (PLoS Med 2008; 5: e193) and results from a unique cohort of women exposed in utero to a synthetic estrogen, diethylstilbestrol, indicated that this exposure increased the women's risk for breast cancer decades later (Cancer Epidemiol Biomarkers Prev. 2006;15:1509-14).

In the context of this Innovator project, several important findings were generated from all five component projects (CPs). We found that not only higher birth weight, but also higher velocity of immediate postnatal growth is a predictor of increased breast cancer risk, notably among premenopausal women (CP1 – Lagiou et al. Br J Cancer. 2008;99:1544-8 – see Appendix 2). We also found that the positive association of birth weight with breast cancer risk several decades later is mediated, at least in part, through higher mammographic density (CP2 – Tamimi et al. Int J Cancer. 2010;126:985-91 – see Appendix 2). We further found suggestive, but intriguing, evidence that genetic predisposition for breast cancer, particularly with respect to SNP rs2981582, may disproportionally increase breast cancer risk among women of higher birth weight (CP3 – Tamimi et al. Cancer Causes Control. 2010;21:689-96 – see Appendix 2).

In the context of CP4, we found evidence that the sharp ecological contrast in breast cancer incidence between Caucasian women in the west and Chinese women in Asia may be due to the much higher levels of umbilical cord levels of sex hormone binding globulin (SHBG) among Chinese compared to Caucasian (CP4 – Lagiou et al. Ann Oncol. 2010 Oct 13. [Epub ahead of print] – see Appendix 2), and / or the fact that umbilical cord levels of IGF-1 are higher and dominate fetal growth among Caucasian, whereas umbilical cord levels of IGF-2 are higher and dominate fetal growth among Chinese women (CP4 – Lagiou et al. Br J Cancer. 2009;100:1794-8 – see Appendix 2). The differential actions of IGF-1 and IGF-2 in embryonic life could be accounted for by the fact that both IGFs are known to bind to the signaling IGF-1 receptor, whereas IGF-2 also binds to the non-signaling IGF-2 receptor. Importantly, the positive associations of IGF-1 with birth size among Caucasian women and of IGF-2 with birth size among Chinese women are almost exclusively evident among women whose higher stature does

not impose physical constrains to fetal growth (CP4 – Lagiou et al. Br J Cancer. 2009;100:1794-8 – see Appendix 2).

The difference in birth size between Caucasian and Asian newborn can account for only a small fraction of the ecological contrast in breast cancer incidence between the two populations. However, the fact that endocrine perinatal influences on birth size are evident almost exclusively among newborn of taller women allows an explanation for the sharp contrast in breast cancer incidence between Caucasian and Asian women and the gradual assimilation of breast cancer incidence among Chinese migrants to the west. Birth size is positively associated with adult height and, over successive generations, improved nutrition, leading to increased adult body size, might reduce constraints on fetal growth and birth size, which in turn affects adult height. The cycle tends to repeat itself, notably over consecutive generations of Asians migrating to the west, and is associated with a gradual increase of breast cancer incidence among them (Overall Innovator project - Lagiou and Trichopoulos. PLoS Med. 2008;5(9):e194 – see Appendix 3).

In the context of CP5, birth weight was found to be significantly positively associated with hematopoietic stem cell measurements among term newborns, supporting a role of stem cell pool on cancer risk (CP5 – Strohsnitter et al. Br J Cancer. 2008;98:660-3 – see Appendix 2). We then analyzed mononuclear cells derived from human umbilical cord blood for gene expressions of markers reported for putative breast epithelial stem cells and progenitors and quantified such cell populations by flow cytometry (CP5 – Qiu et al. Ann Oncol. 2011; in press – see Appendix 2). Thus, we were able to document the existence of putative breast stem/progenitor cell phenotypes in umbilical cord blood - to our knowledge, this is the first report of measurable breast stem cells in human umbilical cord blood.
We also found that umbilical cord blood concentrations of certain breast stem/progenitor cell phenotypes correlate positively with those of hematopoietic stem cells (CP5 – Qiu et al. Ann Oncol. 2011; in press – see Appendix 2). The studies in the context of CP5 provide a unique insight in the biological processes underlying the early life roots of breast cancer, and indicate that the size of the mammary tissue specific stem cell pool and the intrauterine factors that affect it may determine breast cancer risk in adult life.

So what?

Etiological research, a prerequisite for the primary prevention of breast cancer, has so far concentrated in adult life and the results leave much to be desired. Evidence generated in the context of the present Innovator project indicates that important causal factors could be operating much earlier, as early as *in utero.* More importantly, our results point to specific biological processes that appear to underlie the early life roots of breast cancer.

On the basis of results generated and insights gained in the context of the W81XWH-05-1-0314 Innovator project, new avenues for research can be envisaged. It would be important to examine whether the balance of IGF-1 and IGF-2 in the cord blood underlies the lower breast cancer risk of women born after pregnancy toxemia, in comparison to women born after a normal pregnancy. It would also be important to examine whether cord blood IGF-2 is less closely linked to the size of the embryonic stem cell pool compared to cord blood IGF-1, since the size

of this pool is associated with birth weight, which in turn is associated with mammographic density, a strong predictor of breast cancer risk.

With a view to identify possible preventive interventions, future studies could focus on factors that may affect cord blood levels of SHBG, IGF-1 and IGF-2, including maternal diet and other maternal characteristics or lifestyle habits. They could also investigate whether cord blood IGF-1 and IGF-2 levels are associated with epigenetic processes, that is, biochemical changes in the genes which are heritable but reversible, in the umbilical cord blood stem cell populations to further elucidate mechanisms of a woman's susceptibility to future risk of breast cancer.

It may not be realistic to expect eradication of breast cancer in the foreseeable future. We know, however, that Caucasian women in the west have 4 times higher the breast cancer risk of Chinese women in Asia. Adult life risk factors between the two populations cannot explain but a fraction of this striking difference. Moreover, this difference in incidence cannot be accounted for by genetic factors, since breast cancer risk among Chinese migrants to the US rises over several generations and approaches that among US Caucasian. This is not to say that genetic factors are not important in breast cancer occurrence - indeed, we know that they are. However, the differences in the prevalence of the responsible genes between Caucasian and Chinese are too small to account for anything but a very small fraction of this striking difference. Thus, we believe that a reduction of breast cancer incidence in the western world by up to 50% and among contemporary Chinese by up to 20% over the next few generations would be achievable if we understood the perinatal factors responsible for the difference in breast cancer incidence between Chinese in Asia and Caucasian in the west. The focus on perinatal factors is justified by the fact that, among first generation Chinese migrants to the US, the incidence of breast cancer changes little, which confirms that genes and adult life breast cancer risk factors play secondary roles in the ecological contrast between the two populations.

REFERENCES (papers supported by the W81XWH-05-1-0314 Innovator Award)

Publications addressing the issues posed in the five component projects of the W81XWH-05-1-0314 Innovator Award [Copies in Appendix 2]

(CP1)
Lagiou P, Hsieh CC, Trichopoulos D, Adami HO, Hall P, Chie L, Ekbom A. Neonatal growth and breast cancer risk in adulthood. Br J Cancer. 2008;99:1544-8.

(CP2)
Tamimi RM, Eriksson L, Lagiou P, Czene K, Ekbom A, Hsieh CC, Adami HO, Trichopoulos D, Hall P. Birth weight and mammographic density among postmenopausal women in Sweden. Int J Cancer. 2010;126:985-91.

(CP3)

Tamimi RM, Lagiou P, Czene K, Liu J, Ekbom A, Hsieh CC, Adami HO, Trichopoulos D, Hall P. Birth weight, breast cancer susceptibility loci, and breast cancer risk. Cancer Causes Control. 2010;21:689-96.

(CP4)
Lagiou P, Hsieh CC, Lipworth L, Samoli E, Okulicz W, Troisi R, Xu B, Hall P, Ekbom A, Adami HO, Trichopoulos D. Insulin-like growth factor levels in cord blood, birth weight and breast cancer risk. Br J Cancer. 2009;100:1794-8.

Lagiou P, Samoli E, Okulicz W, Xu B, Lagiou A, Lipworth L, Georgila C, Vatten L, Adami HO, Trichopoulos D, Hsieh CC. Maternal and cord blood hormone levels in the United States and China and the intrauterine origin of breast cancer. Ann Oncol. 2010 Oct 13. [Epub ahead of print]

(CP5)
Strohsnitter WC, Savarese TM, Low HP, Chelmow DP, Lagiou P, Lambe M, Edmiston K, Liu Q, Baik I, Noller KL. Adami H-O, Trichopoulos D, Hsieh C-C. Correlation of umbilical cord blood haematopoietic stem and progenitor cell levels with birth weight: implications for a prenatal influence on cancer risk. Br J Cancer. 2008;98:660-3.

Qiu L, Low HP, Chang CI, Strohsnitter WC, Anderson M, Edmiston K, Adami HO, Ekbom A, Hall P, Lagiou P, Trichopoulos D, Hsieh CC. Novel measurements of mammary stem cells in human umbilical cord blood as prospective predictors of breast cancer susceptibility in later life. Ann Oncol. 2011; in press.

Perspectives / commentaries on the early life origins of breast cancer supported by the W81XWH-05-1-0314 Innovator Award (activities in the context of CP6) [Copies in Appendix 3]

Lagiou P, Trichopoulos D. Birth size and the pathogenesis of breast cancer. PLoS Med. 2008;5(9):e194.

Trichopoulos D, Adami HO, Ekbom A, Hsieh CC, Lagiou P. Early life events and conditions and breast cancer risk: From epidemiology to etiology. Int J Cancer. 2008;122:481-485.

Other publications on indirectly relevant research questions with reference to support by the W81XWH-05-1-0314 Innovator Award [Copies in Appendix 4]

Faupel-Badger JM, Hsieh CC, Troisi R, Lagiou P, Potischman N. Plasma volume expansion in pregnancy: implications for biomarkers in population studies. Cancer Epidemiol Biomarkers Prev. 2007;16:1720-3.

Li J, Eriksson L, Humphreys K, Czene K, Liu J, Tamimi R, Lindstrom S, Hunter DJ, Vachon C, Couch F, Christopher S, Lagiou P, Hall P. Genetic variation in the estrogen metabolic pathway and mammographic density as an intermediate phenotype of breast cancer. Breast Cancer Res. 2010 Mar 9;12(2):R19. [Epub ahead of print]

Troisi R, Lagiou P, Trichopoulos D, Xu B, Chie L, Stanczyk FZ, Potischman N, Adami HO, Hoover RN, Hsieh CC. Cord serum estrogens, androgens, insulin-like growth factor-I, and insulin-like growth factor binding protein-3 in Chinese and U.S. Caucasian neonates. Cancer Epidemiol Biomarkers Prev. 2008;17:224-31.

Lagiou P, Samoli E, Lipworth L, Lagiou A, Fang F, Rossi M, Xu B, Yu GP, Adami HO, Hsieh CC, Trichopoulos D. Energy intake during pregnancy in relation to offspring gender by maternal height. Eur J Epidemiol. 2011;26:39-44.

APPENDIX 1

Personnel paid by the W81XWH-05-1-0314 Innovator award

Core research group

Dimitrios Trichopoulos, PI
Pagona Lagiou, project coordinator
Anders Ekbom
Per Hall
Chung-Cheng Hsieh
Hans-Olov Adami

Coordinating administrators of the overall Innovator Award

Rosa Veras – Grant Administrator
Aida Zejnullahu – Administrative Assistant

Other scientific personnel

Fredrik Granath - statistician
Charlotte Stenson - research assistant
Mattias Söderberg - research assistant
Tobias Svensson - statistician
Ulf Berggren - programmer/database manager
Hatef Darabi - biostatistician
Lena Rosenberg – PhD student/post-doc
Louise Eriksson – PhD student
Todd Savarese – senior researcher
Hoi Pang Low – researcher
Qin Liu – researcher
Li Qiu – researcher

Administrators and technical personnel

Ann Almqvist, senior administrator
Anci Adestam administrator
Gerd Agerberg, data abstractor
Ann-Sofie Andersson, data abstractor
Milka Krestelica, data abstractor
Agneta Lönn, data abstractor
Margareta Rodensjö, technical assistant

APPENDIX 2

PDF copies of publications addressing the issues posed in the five component projects of the W81XWH-05-1-0314 Innovator Award

British Journal of Cancer (2008), 1 5

www.bjcancer.com

Full Paper

Neonatal growth and breast cancer risk in adulthood

P Lagiou[1,2,3], C-C Hsieh[1,3,4], D Trichopoulos[*,1,3], H-O Adami[1,3], P Hall[3], L Chie[5] and A Ekbom[6]

[1]Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA; [2]Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, University of Athens, 75 M Asias Street, Athens GR 115 27, Greece; [3]Department of Medical Epidemiology and Biostatistics, Box 281, Karolinska Institutet, Stockholm SE 171 77, Sweden; [4]Division of Biostatistics and Epidemiology, UMass Cancer Center, University of Massachusetts Medical School, 364 Plantation Street, LRB 427, Worcester, MA 01605,USA; [5]Department of Obstetrics and Gynecology, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215, USA; [6]Clinical Epidemiology Unit, Department of Medicine, Karolinska Institutet, Karolinska University Hospital, Stockholm SE 171 76, Sweden

Birth size of a woman has been positively associated with her breast cancer risk, particularly before menopause, but no study has investigated neonatal growth in relation to this risk. We conducted a case control study nested within a population based cohort of women, born in Sweden between 1901 and 1961, covering all 405 breast cancer patients and 1081 age and hospital matched controls, who were born after newborn charts became available. Compared to neonates who lost <200 g after birth and grew at a rate <25 g day^{-1} after reaching postnatal weight nadir (ie, the minimum, before starting to regain weight), those who either lost $\geqslant 200$ g after birth or grew $\geqslant 25$ g day^{-1} after nadir, or both, were at an approximately 50% increased breast cancer risk. The excess risk was striking and statistically significant among women below 50 years of age, but was not evident among older women. Immediate postnatal weight loss (an indicator of water loss, likely to reflect water retention associated with pregnancy hormones) as well as neonatal weight gain rate after the nadir (known to reflect growth hormone levels) was significantly positively associated with premenopausal breast cancer risk.

British Journal of Cancer advance online publication, 30 September 2008; doi:10.1038/sj.bjc.6604702 www.bjcancer.com

Keywords: breast cancer; postnatal growth; birth weight; early life; perinatal

There is much evidence that birth size of women influences their breast cancer risk (Michels and Xue, 2006; Park *et al*, 2008), particularly before menopause (World Cancer Research Fund/ American Institute for Cancer Research, 2007). No study, however, has investigated neonatal growth in relation to breast cancer risk, even though neonatal growth could be of particular importance, as it is strongly associated with neonatal IGF 1 levels (Albertsson Wikland *et al*, 1998; Ogilvy Stuart *et al*, 1998; Hikino *et al*, 2001; Skalkidou *et al,* 2003). IGF 1 levels, which could track through life, have been associated with breast cancer risk, particularly premenopausal breast cancer risk (Renehan *et al*, 2004; Fletcher *et al*, 2005; Rinaldi *et al*, 2006).

Evaluating neonatal growth is complicated because weight declines during the first few days after birth, mostly because of water loss, before starting to increase (Macdonald *et al*, 2003). The decline is likely to reflect the extent of water retention by the newborn at the time of delivery, under the influence of pregnancy hormones, including oestrogens (Stachenfeld and Keefe, 2002; Gomella *et al*, 2004; Stachenfeld and Taylor, 2004). The rate of weight gain after the nadir is influenced by growth factors, notably the IGF system and its determinants (Albertsson Wikland *et al*, 1998; Ogilvy Stuart *et al*, 1998; Hikino *et al*, 2001; Skalkidou *et al* 2003).

We have investigated neonatal growth in relation to breast cancer in adult life by a case control study nested within a population based cohort of Swedish women.

MATERIALS AND METHODS

Participants

In Sweden, all residents have equal access to the governmental health care system, and because there is essentially no private in patient treatment, hospital services are population based. Moreover, since 1 January 1947, all residents are assigned an individually unique nine digit national registration number, which contains information on the date of birth and the county in which the individual resided in 1947 or the county of birth for those born in 1947 or later. This number allows linkage with several Swedish registries, including the Swedish National Cancer Registry (Lunde *et al*, 1980).

In the mid 1990s, we studied the intrauterine environment in relation to breast cancer risk in the offspring using information from a cohort of women who had been born in one of the five participating hospitals in the Uppsala Örebro Health Care Region from 1874 through 1961 and who had survived at least until 1 January 1958, when the Swedish National Cancer Registry was established (Ekbom *et al*, 1997). In that study, a total of 1068 cases

*Correspondence: Dr D Trichopoulos; E mail: dtrichop@hsph.harvard.edu
Received 25 July 2008; revised 5 September 2008; accepted 10 September 2008

were diagnosed until 1994 and 2727 controls were included (Ekbom *et al*, 1997).

For this study, we retrieved newborn charts with information on postnatal growth of neonates until their discharge. The maternity wards in the five hospitals started to use newborn charts at different calendar periods, and so newborn charts were available for 406 of the 1068 eligible case patients and for 1083 of the 2727 eligible controls, all born from 1901 onwards. Because extreme prematurity has been associated with breast cancer risk (Ekbom *et al*, 2000), we excluded babies born before 32 weeks of gestation (one case and two controls), leaving 405 cases and 1081 controls. Of the former, 90 were below the age of 40 years, 168 were aged 40 49 years, 119 were aged 50 59 years, whereas 28 were aged 60 68 years. The corresponding numbers among controls were 245, 485, 290, and 61 women. In our sample, older women are underrepresented among cases because the National Cancer Registry began in 1958, when many older women belonged to cohorts born before the linked neonatal records became available. The ratio of controls to cases is lower among women 50 years of age or above at breast cancer diagnosis. Thus, among women below the age of 50 years, the control to case ratio is 2.8 (730/258), whereas among older women it is 2.4 (351/147). This is because in the earlier years, when older women were born, the likelihood of recording weight changes of newborns was much lower (when an index case was removed because of missing records, the corresponding controls were also removed, whereas if one or two controls had missing records, the remaining control(s) would suffice for retaining the corresponding case in the analysis).

At the birth of our subjects, breastfeeding predominated for newborns and the mother and child were usually discharged when the baby reached its birth weight. Generally, newborns lose weight during the first week and then gain weight (Macdonald *et al*, 2003). To examine whether these two different phases of postnatal pattern of growth were associated with subsequent risk of breast cancer, we determined maximum postnatal weight loss (defined as (birth weight) (the lowest weight in the hospital)) and the rate of growth since the nadir (defined as (weight at discharge weight at nadir)/(day at discharge day at nadir)).

On the basis of the literature (Macdonald *et al*, 2003) we have created the following five mutually exclusive categories: (a) neonates who remained at the maternity wards for more than 21 days without regaining their birth weight these neonates were analysed separately because their weight loss and gain were unusual; (b) neonates with a maximum weight loss of <200 g and growth rate after nadir <25 g day^{-1}; (c) neonates with a maximum weight loss of $\geqslant 200$ g and growth rate after nadir <25 g day^{-1}; (d) neonates with a maximum weight loss of <200 g and growth rate after nadir $\geqslant 25$ g day^{-1} and (e) neonates with a maximum postnatal weight loss of $\geqslant 200$ g and growth rate after nadir $\geqslant 25$ g day^{-1}. All neonates in categories b e remained at the maternity wards for a maximum of 21 days. The weight loss cutoff of 200 g was a round figure derived from the 6.6% reported to be the median percent of birth weight loss for breastfed children (Macdonald *et al*, 2003), and so with birth weight around 3000 g, we have 3000 g*0.066 $\simeq$ 200 g. The cutoff for the daily rate of growth after nadir was rounded at 25 g day^{-1}, as the reported median time for birth weight recovery among breastfed children is 8.3 days (Macdonald *et al*, 2003), so that 200 g divided by 8.3 days equals approximately 25 g day^{-1}.

Statistical analyses

The statistical analyses were undertaken by modelling the data through conditional logistic regression using PROC PHREG of the SAS statistical software (version 9, SAS Institute, Cary, NC, USA). Covariates adjusted in the analysis included maternal age (in years as a continuous variable), maternal socioeconomic status (low, medium, and high as an ordinal variable), maternal parity (1, 2, and $\geqslant 3$ as categorical indicator variables), pregnancy toxaemia (yes/no), neonatal jaundice (yes/no), twin membership (singleton, monozygotic, and dizygotic as categorical indicator variables), and birth weight (<2500, 2500 2999, 3000 3499, 3500 3999, and $\geqslant 4000$ g as categorical indicator variables).

The study was approved by the Institutional Review Boards of the Karolinska Institutet, Sweden, the Harvard School of Public Health, USA, and the US Department of Defense.

RESULTS

Table 1 presents the maternal and perinatal characteristics of breast cancer patients and their control women (matched to the cases with variable matching ratio). As reported earlier (Ekbom *et al*, 1997), neonatal jaundice is more common among cases, whereas maternal toxaemia is more common among controls. In this data set, the association between birth size and breast cancer risk was weak and statistically non significant (Ekbom *et al*, 1997). Spearman's correlation coefficients of birth weight with maximum weight loss and daily weight gain since nadir were 0.48 ($P<0.0001$) and 0.02 ($P=0.55$), respectively. In these bivariate and possibly confounded patterns, neonatal weight loss appears more pronounced among cases than among controls. There is also some evidence that weight gain after nadir is more pronounced among breast cancer patients below the age of 50 years compared with controls.

After controlling for confounding through conditional logistic regression, we found no evidence that neonates who did not conform to the usual growth pattern are at different breast cancer risk when compared with the reference category of neonates who lost less than 200 g after birth and grew at a rate less than 25 g day^{-1} after nadir (Table 2). In contrast, however, neonates who lost $\geqslant 200$ g after birth, or neonates who grew at a rate of $\geqslant 25$ g day^{-1} after nadir, or neonates with both of these growth pattern characteristics were at an approximately 50% increased risk in later life when compared with the reference category. The excess risk was evident and statistically significant exclusively among women below the age of 50 years, who were presumably premenopausal at breast cancer diagnosis. As, in our data, women were designated as pre or postmenopausal relying only on their age, we have evaluated whether there is an interaction between age and growth pattern with respect to breast cancer risk, and the results were of borderline significance ($P\sim 0.06$).

DISCUSSION

In our case control study, nested within a well defined population based cohort of Swedish women, we have found evidence that immediate postnatal weight loss of the newborn, as well as the neonate's weight gain rate after reaching a nadir of postnatal weight, are significantly positively associated with breast cancer risk among women below the age of 50 years. As indicated in the Introduction, in the light of the literature, we considered the immediate postnatal weight loss as an indicator of water loss, probably reflecting water retention caused by pregnancy hormones, and the postnadir rate of growth as an indicator of higher levels of growth hormones, particularly IGF 1.

We interpret our findings as indicating that higher levels of pregnancy hormones and growth hormones during the immediate postnatal period, particularly IGF 1, play an important role in premenopausal breast cancer risk several decades later.

No association of postnatal growth pattern with breast cancer risk was evident among women 50 years of age or above, and presumably postmenopausal who, however, were relatively few in this study sample. Besides the numbers, it is also possible that any effect of perinatal factors on risk is gradually diluted as additional adult life breast cancer risk factors are introduced, in line with the

Table 1 Maternal and perinatal characteristics of 405 breast cancer cases and 1081 matched control subjects

	All women				Women <50 years old				Women ⩾50 years old			
	Cases		Controls		Cases		Controls		Cases		Controls	
	N 405		*N* 1081		*N* 258		*N* 730		*N* 147		*N* 351	
	No	(%)	No	(%)	No	(%)	No	(%)	No	(%)	No	(%)
Maternal age (years)												
<24	125	30.9	385	35.6	77	19.0	257	23.8	48	11.9	128	11.8
25 34	213	52.6	524	48.5	138	34.1	345	31.9	75	18.5	179	16.6
35+	67	16.5	172	15.9	43	10.6	128	11.8	24	5.9	44	4.1
Maternal socioeconomic status												
Low	303	74.8	804	74.4	190	46.9	543	50.2	113	27.9	261	24.1
Medium	91	22.5	199	18.4	61	15.1	132	12.2	30	7.4	67	6.2
High	11	2.7	78	7.2	7	1.7	55	5.1	4	1.0	23	2.1
Maternal parity												
1	170	42.0	499	46.2	110	27.2	343	31.7	60	14.8	156	14.4
2	107	26.4	283	26.2	70	17.3	204	18.9	37	9.1	79	7.3
3+	128	31.6	299	27.7	78	19.3	183	16.9	50	12.3	116	10.7
Maternal toxaemia												
No	397	98.0	1036	95.8	251	62.0	703	65.0	146	36.0	333	30.8
Yes	8	2.0	45	4.2	7	1.7	27	2.5	1	0.2	18	1.7
Neonatal jaundice												
No	380	93.8	1045	96.7	248	61.2	716	66.2	132	32.6	329	30.4
Yes	25	6.2	36	3.3	10	2.5	14	1.3	15	3.7	22	2.0
Twin membership												
No	398	98.3	1068	98.8	253	62.5	723	66.9	145	35.8	345	31.9
Yes	7	1.7	13	1.2	5	1.2	7	0.6	2	0.5	6	0.6
Birth weight (g)												
<2500	14	3.5	42	3.9	8	2.0	26	2.4	6	1.5	16	1.5
2500 2999	57	14.1	157	14.5	36	8.9	108	10.0	21	5.2	49	4.5
3000 3499	145	35.8	408	37.7	85	21.0	279	25.8	60	14.8	129	11.9
3500 3999	143	35.3	338	31.3	101	24.9	224	20.7	42	10.4	114	10.5
⩾4000	46	11.4	136	12.6	28	6.9	93	8.6	18	4.4	43	4.0
Hospital stay ⩾21 days												
No	372	91.9	1018	94.2	244	60.2	692	64.0	128	31.6	326	30.2
Yes	33	8.1	63	5.8	14	3.5	38	3.5	19	4.7	25	2.3
Maximum weight loss (g) after delivery (among normal discharge)												
<200	73	19.6	247	24.3	48	19.7	162	23.4	25	19.5	85	26.1
⩾200	299	80.4	771	75.7	196	80.3	530	76.6	103	80.5	241	73.9
Weight gain (g day $^{-1}$) after reaching nadir (among normal discharge)												
<25	183	49.2	512	50.3	119	48.8	371	53.6	64	50.0	141	43.3
⩾25	189	50.8	506	49.7	125	51.2	321	46.4	64	50.0	185	56.7
Weight change after delivery (combining previous two variables)												
<200 g/<25 g day $^{-1}$	33	8.9	131	12.9	19	7.8	96	13.9	14	10.9	35	10.7
⩾200 g/<25 g day $^{-1}$	150	40.3	381	37.4	100	41.0	275	39.7	50	39.1	106	32.5
<200 g/⩾25 g day $^{-1}$	40	10.8	116	11.4	29	11.9	66	9.5	11	8.6	50	15.3
⩾200 g/⩾25 g day $^{-1}$	149	40.1	390	38.3	96	39.3	255	36.8	53	41.4	135	41.4

conclusion of a major review that birth weight is positively associated with breast cancer risk mostly among premenopausal women (World Cancer Research Fund/American Institute for Cancer Research, 2007). Age at and type of menopause (natural or induced) are important postmenopausal risk factors, and pre and postmenopausal breast cancer are frequently treated as distinct entities in studies focusing on their hormonal and non hormonal aetiology (Hankinson *et al*, 2008).

Our study makes use of the unusual opportunities available in Sweden for linking population based databases and registries. The nested case control study design retains the advantages of a cohort study in terms of minimisation of information and selection bias. Exclusions were simply on the basis of the availability of linked newborn charts. The sample contained many more women below the age of 50 years (presumably premeno pausal) than older women (presumably postmenopausal), and

Table 2 Conditional logistic regression derived[a] odds ratios (ORs) and 95% confidence intervals (CIs) for breast cancer in relation to the patterns of postnatal weight change

	All women			Women <50 years old (presumably premenopausal)			Women ⩾50 years old (presumably postmenopausal)		
	OR	**95% CI**	***P*-value**	**OR**	**95% CI**	***P*-value**	**OR**	**95% CI**	***P*-value**
Maximum weight loss/ daily weight gain since nadir									
<200 g/<25 g day	Reference			Reference			Reference		
⩾200 g/<25 g day	1.53	0.96 2.44	0.08	1.81	1.00 3.25	0.048	0.95	0.41 2.19	0.90
<200 g/⩾25 g day	1.39	0.78 2.45	0.26	2.33	1.13 4.78	0.02	0.50	0.18 1.39	0.19
⩾200 g/⩾25 g day	1.57	0.98 2.51	0.06	2.04	1.12 3.74	0.02	0.91	0.40 2.03	0.81
Irregulars (hospitalised for >21 days)	1.37	0.64 2.90	0.42	1.09	0.40 2.95	0.87	1.35	0.39 4.67	0.64

[a]Controlling for maternal age, maternal socioeconomic status, maternal parity, pregnancy toxaemia, neonatal jaundice, twin membership, and birth weight.

so there should be more confidence in the associations found among the former than on their absence among presumably postmenopausal women. In the base of the study, on which we relied, birth size indicators (birth weight, birth length, and placental weight) were very weakly positively related to risk, although mutual adjustment of these indicators tended to increase the positive trends (Ekbom *et al*, 1997). However, when a true but weak association is investigated in many studies, some are bound to generate non significant or even null results (Michels and Xue, 2006; Park *et al*, 2008). We had no information about adult life risk factors for breast cancer (e.g., age at menarche), but even if associations of such factors with postnatal growth were to be found, they would probably have been placed as intermediates (which should not be controlled for) rather than as confounders (which should). There are, of course, several other risk factors (e.g., age at the first pregnancy, parity, hormone replacement therapy, and so on), which could not act as confounders, as they are unlikely to be related to postnatal growth.

It has been postulated that the likelihood of breast cancer depends on the number of mammary stem cells, which is determined in early, including intrauterine life, as well as on the early postnatal levels of growth enhancing mammotropic hormones, which affect the replication rate of such stem cells (Trichopoulos, 1990; Adami *et al*, 1995; Trichopoulos *et al*, 2005, 2008). Birth size is known to influence breast cancer risk (Michels and Xue, 2006; Park *et al*, 2008), and there is compelling evidence that periadolescent growth (Ahlgren *et al,* 2004) and adult height (Tretli 1989; World Cancer Research Fund/American Institute for Cancer Research, 2007) are also associated with this risk. Using haematopoietic stem cells as probable correlates of the difficult to measure mammary stem cells, the size of their pool was positively associated with both umbilical cord growth hormones and birth weight (Savarese *et al*, 2007; Strohsnitter *et al*, 2008). No earlier investigation, however, has examined postnatal growth in relation to breast cancer risk, even though postnatal growth is rapid and the mammary gland is far from being fully differentiated (Russo and Russo, 1987).

The IGF system is associated with both breast cancer risk (Renehan *et al*, 2004; Fletcher *et al*, 2005; Rinaldi *et al*, 2006) and postnadir growth (Albertsson Wikland *et al*, 1998; Ogilvy Stuart *et al*, 1998; Hikino *et al*, 2001; Skalkidou *et al,* 2003), and could therefore plausibly explain the association of postnadir growth with this risk. Our explanation of the association of immediate postnatal weight reduction with breast cancer risk invokes higher levels of pregnancy hormones, including oestrogens, on the basis of well known properties of these hormones (Stachenfeld and Keefe, 2002; Gomella *et al*, 2004; Stachenfeld and Taylor, 2004).

Replication of our results is clearly necessary. The examination of the possible differential association of neonatal growth with hormone sensitive and hormone insensitive breast cancer, as reflected for instance in hormone receptor expression (Duffy, 2006; Hankinson *et al*, 2008), would also be of importance. Such information was not available in our database. Animal models have provided valuable information with respect to early life exposures and breast cancer risk (Hilakivi Clarke *et al*, 1994; Hilakivi Clarke and de Assis, 2006) and could be useful in relation to postnatal growth.

The findings of this study are intriguing and the apparent magnitude of effect (the twofold increases in premenopausal breast cancer risk for essentially dichotomous contrasts) indicates that the phenomenon is of considerable importance. Confidence limits, however, are wide and the absence of evidence for even additive interaction is of some concern.

ACKNOWLEDGEMENTS

The study was funded by the W81XWH 05 1 0314 Innovator Award, US Department of Defense Breast Cancer Research Program, and the Office of the Congressionally Directed Medical Research Programs.

REFERENCES

Adami H O, Persson I, Ekbom A, Wolk A, Pontén J, Trichopoulos D (1995) The aetiology and pathogenesis of human breast cancer. *Mutat Res* **333:** 29 35

Ahlgren M, Melbye M, Wohlfahrt J, Sørensen TI (2004) Growth patterns and the risk of breast cancer in women. *N Engl J Med* **351:** 1619 1626

Albertsson Wikland K, Boguszewski M, Karlberg J (1998) Children born small for gestational age: postnatal growth and hormonal status. *Horm Res* **49**(Suppl 2)**:** 7 13

Duffy MJ (2006) Estrogen receptors: role in breast cancer. *Crit Rev Clin Lab Sci* **43:** 325 347

Ekbom A, Erlandsson G, Hsieh C C, Trichopoulos D, Adami H O, Cnattingius S (2000) Risk of breast cancer in prematurely born women. *J Natl Cancer Inst* **92:** 840 841

Ekbom A, Hsieh C C, Lipworth L, Adami H O, Trichopoulos D (1997) Intrauterine environment and breast cancer risk in women. *J Natl Cancer Inst* **88:** 71 76

Fletcher O, Gibson L, Johnson N, Altmann DR, Holly JM, Ashworth A, Peto J, Silva IS (2005) Polymorphisms and circulating levels in the insulin like growth factor system and risk of breast cancer: a systematic review. *Cancer Epidemiol Biomarkers Prev* **14:** 2 19

Gomella TL, Cunningham MD, Eyal FG, Zenk KE (eds) (2004) Neonatology: management, procedures, on call problems, diseases, and drugs. *Nutritional Management*, 5th edn. Chapter 8, p 77. New York, NY: McGraw Hill

Hankinson S, Tamimi R, Hunter D (2008) Breast cancer. In *Textbook of Cancer Epidemiology*, Adami HO, Hunter D, Trichopoulos D (eds) 2nd edn, pp 403 445. New York, NY: Oxford University Press

Hikino S, Ihara K, Yamamoto J, Takahata Y, Nakayama H, Kinukawa N, Narazaki Y, Hara T (2001) Physical growth and retinopathy in preterm infants: involvement of IGF I and GH. *Pediatr Res* **50:** 732 736

Hilakivi Clarke L, Clarke R, Lippman ME (1994) Perinatal factors increase breast cancer risk. *Breast Cancer Res Treat* **31:** 273 284

Hilakivi Clarke L, de Assis S (2006) Fetal origins of breast cancer. *Trends Endocrinol Metab* **17:** 340 348

Lunde As, Lundeborg S, Lettenstrom GS, Thuesen L, Huebner J (1980) *The Person Number Systems of Sweden, Norway, Denmark and Israel*. Hyattsville (MD): National Center for Health Statistics

Macdonald PD, Ross SR, Grant L, Young D (2003) Neonatal weight loss in breast and formula fed infants. *Arch Dis Child Fetal Neonatal Ed* **88:** F472 F476

Michels KB, Xue F (2006) Role of birthweight in the etiology of breast cancer. *Int J Cancer* **119:** 2007 2025

Ogilvy Stuart AL, Hands SJ, Adcock CJ, Holly JM, Matthews DR, Mohamed Ali V, Yudkin JS, Wilkinson AR, Dunger DB (1998) Insulin, insulin like growth factor I (IGF I), IGF binding protein 1, growth hormone, and feeding in the newborn. *J Clin Endocrinol Metab* **83:** 3550 3557

Park SK, Kang D, McGlynn KA, Garcia Closas M, Kim Y, Yoo KY, Brinton LA (2008) Intrauterine environments and breast cancer risk: meta analysis and systematic review. *Breast Cancer Res* **10:** R8

Renehan AG, Zwahlen M, Minder C, O'Dwyer ST, Shalet SM, Egger M (2004) Insulin like growth factor (IGF) I, IGF binding protein 3, and cancer risk: systematic review and meta regression analysis. *Lancet* **363:** 1346 1353

Rinaldi S, Peeters PH, Berrino F, Dossus L, Biessy C, Olsen A, Tjonneland A, Overvad K, Clavel Chapelon F, Boutron Ruault MC, Téhard B, Nagel G, Linseisen J, Boeing H, Lahmann PH, Trichopoulou A, Trichopoulos D, Koliva M, Palli D, Panico S, Tumino R, Sacerdote C, van Gils CH, van Noord P, Grobbee DE, Bueno de Mesquita HB, Gonzalez CA, Agudo A, Chirlaque MD, Barricarte A, Larrañaga N, Quiros JR, Bingham S, Khaw KT, Key T, Allen NE, Lukanova A, Slimani N, Saracci R, Riboli E, Kaaks R (2006) IGF I, IGFBP 3 and breast cancer risk in women: The European Prospective Investigation into Cancer and Nutrition (EPIC). *Endocr Relat Cancer* **13:** 593 605

Russo J, Russo IH (1987) Development of the human mammary gland. In *The Mammary Gland*, Neville MC, Daniel CW (eds), pp 67 93. New York: Plenum Press

Savarese TM, Strohsnitter WC, Low HP, Liu Q, Baik I, Okulicz W, Chelmow DP, Lagiou P, Quesenberry PJ, Noller KL, Hsieh CC (2007) Correlation of umbilical cord blood hormones and growth factors with stem cell potential: implications for the prenatal origin of breast cancer hypothesis. *Breast Cancer Res* **9**(3)**:** R29

Skalkidou A, Petridou E, Papathoma E, Salvanos H, Trichopoulos D (2003) Growth velocity during the first postnatal week of life is linked to a spurt of IGF I effect. *Paediatr Perinat Epidemiol* **17:** 281 286

Stachenfeld NS, Keefe DL (2002) Estrogen effects on osmotic regulation of AVP and fluid balance. *Am J Physiol Endocrinol Metab* **283:** E711 E721

Stachenfeld NS, Taylor HS (2004) Effects of estrogen and progesterone administration on extracellular fluid. *J Appl Physiol* **96:** 1011 1018

Strohsnitter WC, Savarese TM, Low HP, Chelmow DP, Lagiou P, Lambe M, Edmiston K, Liu Q, Baik I, Noller KL, Adami HO, Trichopoulos D, Hsieh CC (2008) Correlation of umbilical cord blood haematopoietic stem and progenitor cell levels with birth weight: implications for a prenatal influence on cancer risk. *Br J Cancer* **98:** 660 663

Tretli S (1989) Height and weight in relation to breast cancer morbidity and mortality. A prospective study of 570 000 women in Norway. *Int J Cancer* **44:** 23 30

Trichopoulos D (1990) Hypothesis: does breast cancer originate in utero? *Lancet* **335:** 939 940

Trichopoulos D, Adami HO, Ekbom A, Hsieh CC, Lagiou P (2008) Early life events and conditions and breast cancer risk: from epidemiology to etiology. *Int J Cancer* **122:** 481 485

Trichopoulos D, Lagiou P, Adami HO (2005) Towards an integrated model for breast cancer etiology: the crucial role of the number of mammary tissue specific stem cells. *Breast Cancer Res* **7:** 13 17

World Cancer Research Fund/American Institute for Cancer Research (2007) *Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective*. Washington DC: AICR

uicc
global cancer control

IJC
International Journal of Cancer

Birth weight and mammographic density among postmenopausal women in Sweden

Rulla M. Tamimi[1,2], Louise Eriksson[3], Pagona Lagiou[1,4], Kamila Czene[3], Anders Ekbom[5], Chung-Cheng Hsieh[1,6], Hans-Olov Adami[1,3], Dimitrios Trichopoulos[1,3] and Per Hall[3]

[1] Department of Epidemiology, Harvard School of Public Health, Boston, MA
[2] Harvard Medical School, Channing Laboratory, Boston, MA
[3] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[4] Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, University of Athens, Athens, Greece
[5] Clinical Epidemiology Unit, Department of Medicine, Karolinska Institutet/Karolinska University Hospital, Stockholm, Sweden
[6] Division of Biostatistics and Epidemiology, UMass Cancer Center, University of Massachusetts Medical School, Worcester, MA

Birth weight is a significant predictor of breast cancer risk in adult life and mammary gland mass could be an intermediate stage in this long process. We have studied the association of birth size measurements with mammographic density, a marker of mammary gland mass. For a population-based sample of 893 postmenopausal women without previous cancer in Sweden, we retrieved information on birth size from birth records and their most recent mammography. Film mammograms of the medio-lateral oblique view were digitized and the Cumulus software was used for computer-assisted semi-automated thresholding of mammographic density. Results were analyzed using generalized linear models controlling for possible confounders. Mean percent mammographic density increased when comparing the extreme categories of birth weight (from 15.6% to 18.6%) and head circumference (from 15.5% to 20.4%), and the corresponding linear trends were statistically significant (*p* values 0.02 and 0.007, respectively). The associations were particularly strong when the cutoff for high versus low mammographic density was set at the relatively high value of 50%. Compared to women weighing 3001–3500 grams at birth, women with birth weights >4000g were at almost 3-fold risk of developing high mammographic density (odds ratio: 2.9, 95% confidence interval 1.1 to 7.9). No association with mammographic density was evident with respect to birth length which, however, is known to be less accurately measured. These results indicate that adult breast density, a powerful predictor of breast cancer risk, has intrauterine roots, as reflected in birth size.

Several investigators have implicitly postulated that perinatal factors are associated with breast cancer risk in the offspring, but the hypothesis that endocrine exposures *in utero* may affect breast cancer risk several decades later was formally articulated in 1990.[1] More than 30 studies have evaluated this hypothesis, most of them focusing on birth weight as an important correlate of intrauterine exposures. Recent metaanalyses have concluded that birth weight is a significant predictor of breast cancer risk in adult life.[2–4]

Key words: birth weight, mammographic density, mammography, breast cancer, birth size
Grant sponsors: This study was funded by the W81XWH 05 1 0314 Innovator Award, US Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs
DOI: 10.1002/ijc.24786
History: Received 22 May 2009; Accepted 16 Jul 2009; Online 29 Jul 2009
Correspondence to: Dimitrios Trichopoulos, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA, E mail: dtrichop@hsph.harvard.edu

The mechanism(s), however, linking intrauterine endocrine exposures and, thus, birth weight to breast cancer risk in later life are not apparent. It has been postulated that mammary gland growth and eventually mass may be a critical manifestation of the early life processes that modulate risk of this disease in adult life.[5] An important test of the validity of the postulated chain of causation linking intrauterine influences to mammary gland mass to breast cancer risk would be through evaluation of the association between birth size and mammary gland mass as assessed through mammographic density. Ever since Wolfe published his first report of an association between breast density and breast cancer risk,[6] studies have confirmed relative risks in the 4 fold range when comparing the highest to the lowest mammographic density categories.[7–9]

Few studies have investigated the association between aspects of birth size and mammographic patterns and, collectively, their results have been inconclusive.[10–14] To our knowledge, however, no study has evaluated this association using at the same time, birth size data from records (as opposed to self recall), and computer assisted mammographic density thresholding methods (as opposed to the more subjective categorical classifications of mammographic density).

In a population based sample of postmenopausal women in Sweden, we have studied the association of aspects of birth

Epidemiology

size as retrieved from birth records with mammographic density evaluated through mammography using the Cumulus software for computer assisted thresholding.

Material and Methods

The study was approved by the Institutional Review Boards of the Karolinska Insitutet, Sweden, the Harvard School of Public Health and the Department of Defense, USA.

Study population

The Cancer and Hormones Replacement in Sweden (CAHRES) study is a nationwide, population based case control study of breast cancer among women 50 to 74 years, born in Sweden and living there at any time between October 1, 1993, and March 31, 1995.[15–17] The current study was based on control subjects within the CAHRES case control study. A total of 4,188 control women were randomly selected from the general population according to the expected age frequency distribution of breast cancer cases (in 5 year age groups) through the Total Population Register (currently the Tax Authority), which also provided their individually unique national registration numbers and addresses. In Sweden, all residents are assigned an individually unique nine digit national registration number, which allows linkage with several Swedish registries. Of the 4,188 controls selected, 3,454 (82%) agreed to participate in the study. Detailed information on reproductive and lifestyle exposures, including use of hormone therapy and oral contraceptives, reproductive and menstrual histories, family history of breast cancer, history of benign breast disease, mammography examinations, self reported anthropometry, occupation, alcohol consumption, and cigarette smoking, were obtained through a mailed questionnaire.

For the present study, we excluded controls with a previous cancer (other than nonmelanoma skin cancer and cancer in situ of the cervix) ($n = 91$), as well as those who were premenopausal ($n = 152$), or had unknown menopausal status ($n = 100$). Thus, there were 3,111 postmenopausal women eligible for this study.

Birth characteristic data

We attempted to identify available birth records for all 3,111 eligible controls in the CAHRES study, using their national registration numbers. We were unable to locate the mother's name in the birth registry for 121 (3.9%) of participants. Using information from the birth registry, research assistants visited 120 hospital and medical record archives to locate the original birth records for the study participants. We were unable to locate birth records or midwife journals with adequate information on birth characteristics for 1243 (41.6%) of the participants with identified mother's name. Because this study included births occurring between 1918 and 1945 throughout all of Sweden, a large proportion of births at this time would have occurred at home, with no recording of birth characteristic information. Thus, in total there were 1,747 (56.2%) women with birth characteristic information. Using a detailed form, we abstracted perinatal characteristic information from available birth records, including birth weight, birth length, and head circumference. Additional information about the birth, including date of birth and twin status, was also recorded.

Mammographic density data

We attempted to collect mammograms from all participants. Using the national registration numbers, we obtained addresses for participants from 1975 to 1995 through the civil registry. During 2007 and 2008, we visited all mammography screening units and radiology departments conducting screening mammography in counties in which participants lived between 1975 and 1995. We obtained film mammograms for 2,042 women who were chosen as controls in the original CAHRES study.

For participants with multiple mammograms, the latest mammogram was selected. The median year of mammogram was 1994 and the interquartile range was 1993–1995, making it unlikely for changes in technology to have introduced substantial variability in the assessment of the outcome under consideration. Film mammograms of the mediolateral oblique view of a randomly selected breast side were digitized using an Array 2905HD Laser Film Digitizer, which covers a range of 0 to 4.7 optical density. The density resolution was set at 12 bit spatial resolution. The Cumulus software used for the computer assisted thresholding was developed at the University of Toronto.[18] For each image, a trained observer (LE) set the appropriate gray scale threshold levels defining the edge of the breast and distinguishing dense from non dense tissue. The software calculated the total number of pixels within the entire region of interest and within the region identified as dense. These values were used to calculate the percentage of the breast area that is dense. A random 10% of the images were included as replicates to assess the intra observer reliability, which was high with a Spearman rank correlation coefficient of 0.95. Mammographic density measurements have been conducted on 1320 women, for 918 of whom birth characteristic information is also available. Because twins are believed to have different *in utero* hormonal exposures compared with singletons, we excluded 8 women who were twin members. In addition, 17 women, who did not have complete covariate information, had to be excluded. Thus, 893 women contributed to the analysis.

Statistical analysis

We used generalized linear models adjusted for covariates to determine the mean percentage mammographic density without transformation, according to birth size categories. To determine whether there was a linear trend of mean percent breast density with increasing birth size, we calculated p values for inclusion of birth size as an ordered (per category) or continuous (per standard deviation SD) variable in the

Table 1. Characteristics of 893 postmenopausal swedish women without breast cancer (recruited as controls in the context of a population based case control study on breast cancer)

	N=893	%
Age at mammogram (years)		
<55	195	21.8
55 59	200	22.4
60 64	174	19.5
65+	324	36.3
Calendar year of mammograms		
Before 1994	244	27.3
1994	302	33.8
1995	345	38.6
1996+	2	0.2
Age at menarche (years)*		
<12	50	6.2
12	127	15.6
13	229	28.2
14	220	27.1
15+	187	23.0
Age at menopause (years)*		
<45	80	10.0
45 49	254	31.7
50 54	377	47.1
55+	90	11.2
Oral contraceptive use*		
Ever	403	45.3
Never	486	54.7
Hormone replacement therapy use*		
Ever	339	38.2
Never	548	61.8
Parity		
0	92	10.3
1	159	17.8
2	340	38.1
3+	302	33.8
Height (cm)		
<160	178	19.9
160 <165	301	33.7
165 <170	273	30.6
170+	141	15.8
Body mass index at recruitment (kg/m^2)		
<20	51	5.7
20 24.9	389	43.6
25 29.9	356	39.9
30.0+	97	10.9
Parity		
0	92	10.3
1	159	17.8
2	340	38.1
3+	302	33.8
Birth weight (g)		
≤2500	32	3.6
2501 ≤3000	145	16.2
3001 ≤3500	324	36.3
3501 ≤4000	291	32.6
>4000	101	11.3
Birth length (cm)*		
<50	232	26.9
50	204	23.7
51	140	16.2
52	130	15.1
53+	156	18.1
Head circumference (cm)*		
<34	170	23.3
34 <35	197	27.0
35 <36	192	26.3
36+	170	23.3

*There were 31 missing values for birth length, 164 for head circumference, 80 for age at menarche, 92 for age at menopause, 4 for oral contraceptive use and 6 for use of hormone replacement therapy use.

model. We included in the multivariate models, as possible confounders, the following known predictors of mammographic density: age at mammogram (continuous), body mass index (BMI) (continuous), parity (none, 1, 2, 3+ births, categorical), and age at menopause (<45, 45 49, 50 54, 55+ years, categorical). Parity and age at menopause have been reported to be associated with mammographic density, but there is little information as to whether they are also associated with birth weight. Nevertheless, we opted for a conservative approach and controlled for them in the analyses. Odds ratios (OR) and 95% confidence intervals (CI) for having high versus low mammographic density also were determined using multiple logistic regression. In this postmenopausal population, we utilized two definitions for high mammographic density: ≥25% and ≥50% breast density. To determine whether there was a linear trend with increasing birth size, we calculated p values from Wald statistics including a continuous term (per SD) in the model. Data analysis was conducted with SAS statistical software version 9.1 (SAS Institute, Cary, NC, USA). All *p* values presented are from two sided tests of statistical significance.

Table 2. Mean percent mammographic density (MD) according to birth size characteristics

	N	Mean MD (%)*	Mean MD (%)†	Mean MD (%)‡
Birth weight (g)				
≤2500	32	14.9	15.6	15.6
2501 ≤3000	145	13.8	13.5	13.6
3001 ≤3500	324	15.6	15.8	15.8
3501 ≤4000	291	18.1	17.8	17.7
>4000	101	17.6	18.3	18.6
P trend**		0.04	0.02	0.02
Birth length (cm)				
<50	232	16.2	16.3	16.4
50	204	16.3	16.1	16.2
51	140	15.6	15.6	15.5
52	130	18.3	18.0	18.1
53+	156	16.5	16.7	16.5
P trend**		0.71	0.83	0.69
Head circumference (cm)				
<34	170	15.8	16.1	15.5
34 <35	197	16.6	16.7	16.9
35 <36	192	16.3	16.1	16.2
36+	170	20.4	20.2	20.4
P trend**		0.04	0.03	0.007

*Adjusted for age (continuous). †Adjusted for age (continuous) and body mass index (continuous). ‡Adjusted for age (continuous), body mass index (continuous), age at menopause (<45, 45 49, 50 54, 54+, missing categorical), parity (0, 1, 2, 3+, categorical). **P* for trend based on per standard deviation increase.

Table 3. Mean area of dense tissue and lucent tissue according to birth size characteristics

	N	Mean dense area (cm^2)*	Mean lucent area (cm^2)*
Birth weight (g)			
≤2500	32	16.3	97.6
2501 ≤3000	145	14.5	104.4
3001 ≤3500	324	16.8	102.1
3501 ≤4000	291	17.8	95.4
>4000	101	20.8	99.5
P trend†		0.02	0.15
Birth length (cm)			
<50	232	16.7	98.2
50	204	17.1	100.0
51	140	16.3	101.7
52	130	18.9	96.8
53+	156	17.4	101.1
P trend†		0.74	0.56
Head circumference (cm)			
<34	170	15.9	98.9
34 <35	197	17.2	99.1
35 <36	192	17.2	99.1
36+	170	21.0	93.8
P trend†		0.01	0.16

*Adjusted for age (continuous), body mass index (continuous), age at menopause (<45, 45 49, 50 54, 54+, missing categorical), parity (0, 1, 2, 3+, categorical). †*P* for trend based on per standard deviation increase.

Results

Among the 893 postmenopausal women, the mean age at the time of the index mammography was 61.2 years, with a SD of 6.8 years. Median percent mammographic density was 11.2 (interquartile range: 4.6 23.3). The mean birth weight of the women was 3452.6 (SD = 506.4) grams, the mean birth length 50.6 (SD = 2.6) cm, and the mean head circumference was 34.4 (SD = 1.5) cm. Frequency distributions by aspects of birth size and possible confounding variables are presented in Table 1 and serve descriptive purposes.

Table 2 shows mean percent mammographic density according to categories of birth size characteristics. The mean values are adjusted first for age at mammogram, then for age at mammogram and BMI, and then for age at mammogram, BMI, age at menopause and parity. Mean percent mammographic density was significantly positively associated with both birth weight and head circumference. In the full models, *p* values for trend were 0.02 and 0.007, respectively. Thus, women with birth weight exceeding 4000 grams had a mean percent mammographic density of 18.6%, while those with birth weight 2500 grams or less had a mean percent mammographic density of 15.6%. Moreover, women with a birth head circumference of 36 cm or more had a mean percent mammographic density of 20.4%, while those with a circumference less than 34 cm had a mean percent mammographic density of 15.5%. Both of these associations were driven by the positive association with dense area on the mammogram (*p* for trend 0.02 for birth weight and 0.01 for head circumference), as there was no association with non dense area on the mammogram (*p* for trend 0.15 for birth weight and 0.16 for head circumference) in multivariate models (Table 3). There were no outliers for birth weight. There were 8 individuals with a head circumference 3 times the interquartile range, but their exclusion did not materially affect the results. We found no association between birth length and percent mammographic density (*p* for trend 0.69 for percent mammographic density, 0.74 for dense area and 0.56 for non dense area on the mammogram).

To further explore the pattern of association of birth size characteristics with mammographic density, we have calculated odds ratios for high versus low mammographic density in relation to birth size characteristics, using, alternatively, 25% and 50% (Table 4) density as cutoffs, and controlling for potentially confounding variables. For all three birth size characteristics,

Epidemiology

Table 4. Odds ratios (OR) and 95% confidence intervals (95% CI) for high versus low mammographic density in relation to birth size characteristics, using 25% and 50% density as cutoff

	Mammographic density			Mammographic density		
Birth weight (g)	High (≥25%)	Low (<25%)	OR (95%CI)[1]	High (≥50%)	Low (<50%)	OR (95%CI)[1]
≤2500	4	28	0.40 (0.12 1.35)	3	29	0.57 (0.18 1.81)
2501 ≤3000	27	118	0.72 (0.41 1.24)	2	143	
3001 ≤3500	75	249	1.0 (REF)	13	311	1.0 (REF)
3501 ≤4000	74	217	0.97 (0.64 1.47)	21	270	1.68 (0.78 3.62)
>4000	21	80	0.98 (0.53 1.79)	8	93	2.91 (1.07 7.88)
P trend[2]			0.31			0.048
Birth length (cm)						
<50	50	182	1.27 (0.70 2.30)	11	221	0.77 (0.27 2.21)
50	50	154	1.65 (0.90 3.01)	9	195	0.85 (0.29 2.52)
51	25	115	1.0 (REF)	7	133	1.0 (REF)
52	37	93	1.65 (0.86 3.14)	9	121	1.10 (0.37 3.26)
53+	34	122	1.22 (0.64 2.33)	10	146	1.04 (0.35 3.11)
P trend[2]			0.48			0.49
Head circumference (cm)						
<34 cm	39	131	0.81 (0.46 1.44)	10	160	0.66 (0.23 1.87)
34 <35 cm	47	150	1.19 (0.70 2.03)	9	188	0.90 (0.33 2.44)
35 <36 cm	43	149	1.0 (REF)	10	182	1.0 (REF)
36+	50	120	1.61 (0.93 2.78)	15	155	1.72 (0.68 4.35)
P trend[2]			0.08			0.04

[1]Adjusted for age (continuous), body mass index (continuous), age at menopause (<45, 45 49, 50 54, 54+, categorical), parity (0, 1, 2, 3+, categorical). [2]P for trend based on per standard deviation increase.

intermediate categories were used as referents. In line with the results shown in Table 2, birth weight and head circumference were positively associated with mammographic density. How ever, the trends were monotonic and statistically significant only when the cutoff was set at 50%, indicating that the associ ation between birth size and mammographic density is not lin ear and is particularly evident at the extreme. Again, in line with the results in Table 2, we found no association of birth length with mammographic density.

There was an age difference between the women consid ered to have high mammographic density (>50%) and women with lower mammographic density. Women with high density were on average 56.6 years of age (interquartile range 51.6 61.0), while women with low density were on av erage 61.5 years of age (interquartile range 56.0 67.0). In all analyses, age was adjusted for, but there was not sufficient power in our study to examine possible interaction between age at mammography and birth size in relation to mammo graphic density.

We have also explored whether the association between birth size and mammographic density might be mediated through measures of growth and development, notably age at menarche and adult height, by alternatively and simultane ously including these variables in the models. Height and, to a lesser extent, age at menarche appear to mediate, in part, the associations of both birth weight and head circumference with mammographic density. Thus, the significant associa tions of mammographic density with these two measures of birth size, after adjustment for age alone (p for trend 0.04 in both instances) became non significant after controlling for both height and age at menarche (p for trend 0.24 for birth weight and 0.17 for head circumference). Percent mammo graphic density was weakly associated with both height (Spearman correlation = +0.09, p value = 0.009) and age at menarche (Spearman correlation = +0.04; p value = 0.31).

Discussion

In our large population based sample of postmenopausal Swedish women, birth weight and head circumference were both significantly positively associated with postmenopausal mammographic density. The 3 5% difference in mammo graphic density comparing the extreme categories of birth weight or head circumference is not trivial and is indeed comparable to that observed between women with and with out hormonal treatment.[19,20] Although birth length was not associated with mammographic density, this could reflect the fact that birth length is measured less reliably than birth weight and head circumference.[21,22] Moreover, we found evi dence that part of the association between birth size and mammographic density could be explained by the association

of birth size with height and, to a lesser extent, with age at menarche.

This is the first study to evaluate birth size parameters retrieved from birth records in relation to mammographic density assessed by a semi automated method. Earlier studies assessed breast density using a more subjective categorical measure of breast density or relied on self reported birth weight. In a study of Swedish women ($n = 370$), high density categories of Wolfe's parenchymal patterns (*i.e.,* P2/DY) were non significantly positively associated with birth weight and significantly with placental weight.[10] In a British cohort of 1,298 perimenopausal women, McCormack and colleagues found no association between birth weight and Wolfe's parenchymal patterns.[11] The Glasgow Alumni Cohort study ($n = 628$) also reported no association between self reported birth weight and a radiologist assessed six category classification of mammographic density.[13,14] In contrast, Cerhan and colleagues found a significant positive association between self reported birth weight and mammographic density among postmenopausal women using the same computer assisted measure of mammographic density as in our study.[12]

Inconsistencies in previous studies might arise due to measurement error of mammographic density and/or birth characteristic data. Computer assisted thresholding methods, which are less subjective than categorical classifications, have been demonstrated to be better predictors of breast cancer risk.[7,23] In addition, birth data from records were significantly associated with breast cancer risk, while those from self recall were not associated, suggesting that there is significant measurement error in self reported measures of birth characteristics.[3]

Mammographic density is one of the strongest known risk factors for breast cancer, with women in the highest category of mammographic density being at a four to sixfold increased risk of breast cancer as compared with women in the lowest category.[6,7,9] Although the biologic mechanism underlying this risk remains unclear, use of exogenous hormones increases mammographic density[24–30] whilst tamoxifen (a selective estrogen receptor modulator with anti estrogenic effects in the breast parenchyma) reduces breast density.[19,31,32] Based on these data, it has been hypothesized that breast density represents cumulative exposure to estrogens.[33] However, studies examining endogenous circulating levels of sex steroids and mammographic density have for the most part found no association.[18,20,34–37]

It is also possible that mammographic density is influenced in part by exposures early in life, including the intrauterine life. It has been shown that birth weight is positively associated with the size of the cord blood stem cell pool, a surrogate for overall stem cell potential.[38] Assessed through mammographic density, mammary gland mass is likely to be correlated with the number of mammary stem cells,[39,40] generating the association of birth weight with mammographic density.

A limitation of the current study is that we were unable to obtain birth characteristic information for all eligible participants. This is due to the fact that the current study was population based and covered all of Sweden in the early 1900s, when many births still look place at home and these characteristics were not recorded. It is highly unlikely, however, that inability to obtain birth records would be associated with adult mammographic density and that selection bias would have been generated. Another shortcoming of our study is that, although all women were postmenopausal at enrolment, the index mammography for a few of them might have been taken prior to the occurrence of menopause. As with any observational study, there is also the potential for confounding. Body mass index is not likely to be an intermediate factor in the association between birth size and mammographic density, as it is strongly affected by energy intake in adolescent and adult life. Since, however, body mass index is a strong predictor of percent mammographic density, it is conceivable that it could exercise confounding influences and for this reason it was finely controlled for in the analyses. With height and weight being self reported, it is possible that measurement error in these variables could lead to residual confounding, even after adjustment for body mass index. An advantage of our study is its reliance on birth records, the information from which is more reliable than that from parental or self recall. Birth records were filled routinely and not in the context of a research protocol and, therefore, differential misclassification with respect to mammographic density is all but impossible. Non differential misclassification is unavoidable, but could only attenuate the reported associations. Another advantage is measurement of mammographic density through a semi automated method, as opposed to earlier more subjective categorical assessments.

In conclusion, our results support the hypothesis that adult breast density, a powerful correlate of breast cancer risk, has intrauterine roots, as reflected in birth size. The results indirectly provide further support to the general hypothesis that breast cancer risk in adulthood is programmed already *in utero* and then modified by a multitude of factors later in life.

Epidemiology

References

1. Trichopoulos D. Hypothesis: Does breast cancer originate in utero? *Lancet* 1990;335: 939 40.
2. Michels KB, Xue F. Role of birthweight in the etiology of breast cancer. *Int J Cancer* 2006;119:2007 25.
3. dos Santos Silva I, De Stavola B, McCormack V, Collaborative Group on Pre Natal Risk Factors and Subsequent Risk of Breast Cancer. Birth size and breast cancer risk: Re analysis of individual participant data from 32 studies. *PLoS Med* 2008;5:e193.
4. Park SK, Kang D, McGlynn KA, Garcia Closas M, Kim Y, Yoo KY, Brinton LA. Intrauterine environments and breast cancer risk: meta analysis and systematic review. *Breast Cancer Res* 2008; 10:R8.

5. Trichopoulos D, Lagiou P, Adami HO. Towards an integrated model for breast cancer etiology: the crucial role of the number of mammary tissue specific stem cells. *Breast Cancer Res* 2005;7:13 7.
6. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer* 1976;37:2486 92.
7. Byrne C, Schairer C, Wolfe J, Parekh N, Salane M, Brinton LA, Hoover R, Haile R. Mammographic features and breast cancer risk: effects with time, age, and menopause status. *J Natl Cancer Inst* 1995;87:1622 9.
8. McCormack VA, Dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta analysis. *Cancer Epidemiol Biomarkers Prev* 2006;15:1159 69.
9. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, Jong RA, Hislop G, Chiarelli A, Minkin S, Yaffe MJ. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007;356:227 36.
10. Ekbom A, Thurfjell E, Hsieh CC, Trichopoulos D, Adami HO. Perinatal characteristics and adult mammographic patterns. *Int J Cancer* 1995;61:177 80.
11. McCormack VA, dos Santos Silva I, De Stavola BL, Perry N, Vinnicombe S, Swerdlow AJ, Hardy R, Kuh D. Life course body size and perimenopausal mammographic parenchymal patterns in the MRC 1946 British birth cohort. *Br J Cancer* 2003;89:852 9.
12. Cerhan JR, Sellers TA, Janney CA, Pankratz VS, Brandt KR, Vachon CM. Prenatal and perinatal correlates of adult mammographic breast density. *Cancer Epidemiol Biomarkers Prev* 2005;14:1502 8.
13. Jeffreys M, Warren R, Gunnell D, McCarron P, Smith GD. Life course breast cancer risk factors and adult breast density (United Kingdom). *Cancer Causes Control* 2004;15:947 55.
14. Jeffreys M, Warren R, Highnam R, Davey Smith G. Breast cancer risk factors and a novel measure of volumetric breast density: cross sectional study. *Br J Cancer* 2008;98: 210 6.
15. Magnusson C, Baron J, Persson I, Wolk A, Bergström R, Trichopoulos D, Adami HO. Body size in different periods of life and breast cancer risk in post menopausal women. *Int J Cancer* 1998;76:29 34.
16. Magnusson CM, Persson IR, Baron JA, Ekbom A, Bergström R, Adami HO. The role of reproductive factors and use of oral contraceptives in the aetiology of breast cancer in women aged 50 to 74 years. *Int J Cancer* 1999;80:231 6.
17. Magnusson C, Baron JA, Correia N, Bergström R, Adami HO, Persson I. Breast cancer risk following long term oestrogen and oestrogen progestin replacement therapy. *Int J Cancer* 1999;81: 339 44.
18. Boyd NF, Stone J, Martin LJ, Jong R, Fishell E, Yaffe M, Hammond G, Minkin S. The association of breast mitogens with mammographic densities. *Br J Cancer* 2002;87:876 82.
19. Brisson J, Brisson B, Coté G, Maunsell E, Bérubé S, Robert J. Tamoxifen and mammographic breast densities. *Cancer Epidemiol Biomarkers Prev* 2000;9:911 5.
20. Greendale GA, Palla SL, Ursin G, Laughlin GA, Crandall C, Pike MC, Reboussin BA. The association of endogenous sex steroids and sex steroid binding proteins with mammographic density: results from the postmenopausal estrogen/progestin interventions mammographic density study. *Am J Epidemiol* 2005;162:826 34.
21. Johnson TS, Engstrom JL, Gelhar DK. Intra and interexaminer reliability of anthropometric measurements of term infants. *J Pediatr Gastroenterol Nutr* 1997; 24:497 505.
22. Johnson TS, Engstrom JL, Haney SL, Mulcrone SL. Reliability of three length measurement techniques in term infants. *Pediatr Nurs* 1999;25:13 7.
23. Brisson J, Diorio C, Masse B. Wolfe's parenchymal pattern and percentage of the breast with mammographic densities: redundant or complementary classifications? *Cancer Epidemiol Biomarkers Prev* 2003;12:728 32.
24. Stomper PC, Van Voorhis BJ, Ravnikar VA, Meyer JE. Mammographic changes associated with postmenopausal hormone replacement therapy: a longitudinal study. *Radiology* 1990;174:487 90.
25. Harvey JA, Pinkerton JV, Herman CR. Short term cessation of hormone replacement therapy and improvement of mammographic specificity. *J Natl Cancer Inst* 1997;89:1623 5.
26. Leung W, Goldberg F, Zee B, Sterns E. Mammographic density in women on postmenopausal hormone replacement therapy. *Surgery* 1997;122:669 73.
27. Persson I, Thurfjell E, Holmberg L. Effect of estrogen and estrogen progestin replacement regimens on mammographic breast parenchymal density. *J Clin Oncol* 1997;15:3201 7.
28. Lundström E, Wilczek B, von Palffy Z, Söderqvist G, von Schoultz B. Mammographic breast density during hormone replacement therapy: differences according to treatment. *Am J Obstet Gynecol* 1999;181:348 52.
29. Vachon CM, Sellers TA, Vierkant RA, Wu FF, Brandt KR. Case control study of increased mammographic breast density response to hormone replacement therapy. *Cancer Epidemiol Biomarkers Prev* 2002;11: 1382 8.
30. Greendale GA, Reboussin BA, Slone S, Wasilauskas C, Pike MC, Ursin G. Postmenopausal hormone therapy and change in mammographic density. *J Natl Cancer Inst* 2003;95:30 7.
31. Chow CK, Venzon D, Jones EC, Premkumar A, O'Shaughnessy J, Zujewski J. Effect of tamoxifen on mammographic density. *Cancer Epidemiol Biomarkers Prev* 2000;9:917 21.
32. Cuzick J, Warwick J, Pinney E, Warren RM, Duffy SW. Tamoxifen and breast density in women at increased risk of breast cancer. *J Natl Cancer Inst* 2004;96: 621 8.
33. Clemons M, Goss P. Estrogen and the risk of breast cancer. *N Engl J Med* 2001;344: 276 85.
34. Aiello EJ, Tworoger SS, Yasui Y, Stanczyk FZ, Potter J, Ulrich CM, Irwin M, McTiernan A. Associations among circulating sex hormones, insulin like growth factor, lipids, and mammographic density in postmenopausal women. *Cancer Epidemiol Biomarkers Prev* 2005;14:1411 7.
35. Tamimi RM, Hankinson SE, Colditz GA, Byrne C. Endogenous sex hormone levels and mammographic density among postmenopausal women. *Cancer Epidemiol Biomarkers Prev* 2005;14:2641 7.
36. Noh JJ, Maskarinec G, Pagano I, Cheung LW, Stanczyk FZ. Mammographic densities and circulating hormones: a cross sectional study in premenopausal women. *Breast* 2006;15:20 8.
37. Warren R, Skinner J, Sala E, Denton E, Dowsett M, Folkerd E, Healey CS, Dunning A, Doody D, Ponder B, Luben RN, Day NE, et al. Associations among mammographic density, circulating sex hormones, and polymorphisms in sex hormone metabolism genes in postmenopausal women. *Cancer Epidemiol Biomarkers Prev* 2006;15:1502 8.
38. Strohsnitter WC, Savarese TM, Low HP, Chelmow DP, Lagiou P, Lambe M, Edmiston K, Liu Q, Baik I, Noller KL, Adami HO, Trichopoulos D, et al. Correlation of umbilical cord blood haematopoietic stem and progenitor cell levels with birth weight: implications for a prenatal influence on cancer risk. *Br J Cancer* 2008;98:660 3.
39. Baik I, Becker PS, DeVito WJ, Lagiou P, Ballen K, Quesenberry PJ, Hsieh CC. Stem cells and prenatal origin of breast cancer. *Cancer Causes Control* 2004;15:517 30.
40. Trichopoulos D, Adami HO, Ekbom A, Hsieh CC, Lagiou P. Early life events and conditions and breast cancer risk: From epidemiology to etiology. *Int J Cancer* 2008;122:481 485.

Cancer Causes Control
DOI 10.1007/s10552 009 9496 7

ORIGINAL PAPER

Birth weight, breast cancer susceptibility loci, and breast cancer risk

Rulla M. Tamimi · Pagona Lagiou · Kamila Czene · Jianjun Liu · Anders Ekbom · Chung-Cheng Hsieh · Hans-Olov Adami · Dimitrios Trichopoulos · Per Hall

Received: 11 November 2009 / Accepted: 22 December 2009

Abstract
Background There is considerable evidence that birth weight is positively associated with breast cancer risk, and seven single-nucleotide polymorphisms (SNPs) have been conclusively associated with this risk. We have hypothesized that breast cancer susceptibility loci may have a greater influence on breast cancer risk among women with higher birth weight, who are expected to have a larger pool of mammary stem cells that are susceptible to malignant transformation.
Patients and methods In the context of a nationwide, population-based case control study in Sweden, we retrieved recorded birth weight for 693 breast cancer cases and 747 control women who were also genotyped for most or all of the seven recently documented breast cancer susceptibility SNPs: rs2981582, rs12443621, rs8051542, rs3803662, rs889312, rs13281615, and rs3817198.
Results We grouped heterozygotes with homozygotes for the wild-type allele, and we found a marginally significant interaction ($p \sim 0.07$) between birth weight and rs2981582 (FGFR2), the genotype repeatedly identified as the top hit in genome-wide association studies. There were similar, though not significant, patterns for the other six SNPs.
Conclusions Although our findings require confirmation, we found suggestive evidence that genetic susceptibility modifies the positive association of birth weight with breast cancer.

Keywords Allele · Birth weight · Breast cancer · Gene · Genotype · Polymorphism

R. M. Tamimi · P. Lagiou · C. C. Hsieh · H. O. Adami · D. Trichopoulos (✉)
Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA
e mail: dtrichop@hsph.harvard.edu

R. M. Tamimi
Channing Laboratory, Harvard Medical School, 181 Longwood Avenue, Boston, MA 02115, USA

P. Lagiou
Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, University of Athens, 75 M. Asias Street, 115 27 Athens, Greece

K. Czene · H. O. Adami · D. Trichopoulos · P. Hall
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, 171 77 Stockholm, Sweden

J. Liu
Population Genetics, Genome Institute of Singapore, Singapore 138672, Singapore

A. Ekbom
Clinical Epidemiology Unit, Department of Medicine, Karolinska Institutet/Karolinska University Hospital, 171 76 Stockholm, Sweden

C. C. Hsieh
Division of Biostatistics and Epidemiology, UMass Cancer Center, University of Massachusetts Medical School, 364 Plantation Street, LRB 427, Worcester, MA 01605, USA

Introduction

Family history of breast cancer is an established risk factor for breast cancer [1]. Twin studies indicate that inherited genetic influences account for approximately 27% of the variation in breast cancer risk [2]. The proportion of breast cancer attributable to rare highly penetrant genes, such as BRCA1 and BRCA2, however, is probably no more than around 5% [3]. Hence, a number of more common variants in lower penetrance genes likely contribute substantially to the familial occurrence of breast cancer and account for

Published online: 07 January 2010

higher attributable rates. Accumulating data support such a polygenic model of inherited breast cancer risk [4], in which each gene may confer only a small amount of risk individually yet in combination, they would result in a noticeable increase in susceptibility to breast cancer [5, 6]. Recently, genome-wide association studies (GWAS) have provided the opportunity to comprehensively search the genome, in an unbiased approach for genetic variants with even modest associations with disease [7]. Through this approach, three GWAS have identified novel breast cancer susceptibility loci [8 10]. The Breast Cancer Association Consortium (BCAC) conducted a multi-stage genome-wide association study of breast cancer and identified seven single-nucleotide polymorphisms (SNPs) associated with breast cancer risk. These variants were genome-wide significant after replication in its final stage of 21,860 cases and 22,578 controls from 22 studies [8]. The Swedish and Singaporean Breast Cancer Association Consortium (SASBAC) contributed to the final replication stage in this consortium.

There is also considerable evidence that birth weight and other aspects of fetal growth are associated with breast cancer risk in later life [11 16]. Although the mechanisms underlying this association are poorly understood, it has been suggested that the size of the mammary stem cell pool and eventually mammary gland mass may represent intermediate manifestations of the early life processes that modulate risk of this disease in adult life [17, 18].

We hypothesized that, if breast cancer risk is positively associated with the size of the pool of mammary stem cells, this association would be stronger in the presence of breast cancer susceptibility loci and weaker or even absent in the absence of such loci. Evidence of an interaction of this type would provide strong support to the hypothesis that the mammary stem cell pool is critical in the intrauterine roots of breast cancer risk in adult life. Using birth weight as the only documented proxy of the stem cell pool [19], we examined the association between birth weight, breast cancer susceptibility loci, and risk of breast cancer in the SASBAC study. To our knowledge, no gene environment interactions have been identified with these novel breast cancer susceptibility loci.

Subjects and methods

Study population

The Swedish and Singaporean Breast Cancer Association Consortium (SASBAC) rely on a Swedish case control study, with genotyping conducted in Singapore. More specifically, the women included are participants in the Cancer and Hormones Replacement in Sweden (CAHRES) study, a nationwide, population-based case control study of incident breast cancer among women 50 74 years old, born in Sweden and living there at any time between October 1, 1993, and March 31, 1995 [20 22]. In this study, breast cancer cases were identified through the six Swedish regional cancer registries in which 98% of all diagnosed cancer cases in Sweden are reported. A total of 3,979 eligible breast cancer cases were identified and invited to participate by their physicians. Of those invited, 3,345 (84%) agreed to participate. Non-participation was due to inability to contact patient or patient refusal in 11% of women and to patient death or physician refusal in the remaining instances. To serve as control subjects, a total of 4,188 women were randomly selected from the general population, according to the expected age frequency distribution of cases (in 5-year age groups), through the Total Population Register (currently the Tax Authority), which also provided their addresses and national registration numbers unique ten-digit numbers assigned to all Swedish residents. Of the 4,188 controls selected in the study, 3,454 (82%) agreed to participate. Detailed information was obtained through a self-administered questionnaire mailed to cases and controls (20 22). The questionnaire covered, among others, reproductive and menstrual histories, family history of breast cancer, benign breast disease, mammography examinations, use of oral contraceptives, use of hormone replacement therapy, current and past anthropometric measurements, profession, alcohol consumption, and cigarette smoking.

Women with a previous cancer (other than non-melanoma skin cancer and cancer in situ of the cervix; $n = 112$ cases and 91 controls), as well as women who were premenopausal ($n = 198$ cases and 152 controls), or had unknown menopausal status ($n = 217$ cases and 100 controls) were excluded. Thus, there were 2,818 cases and 3,111 controls eligible for this study.

The present study was approved by the Institutional Review Boards at Karolinska Institutet, Uppsala University, Harvard School of Public Health and US Department of Defense, and was performed in compliance with the Helsinki Declaration.

Blood sampling

For genotype analysis, 1,500 women with invasive breast cancer and 1,500 controls were originally selected. For reasons unrelated to the objectives of the present investigation, all remaining eligible cases and controls who had taken menopausal hormone treatment for at least 4 years (191 cases and 108 controls), all women with self-reported diabetes mellitus (110 cases and 104 controls), and another 345 controls from the parent study selected for a parallel investigation were also included in the genotype analysis.

Thus, 1,801 women with breast cancer and 2,057 control women were selected for genotype analysis. We contacted the selected women by mail, and those who gave informed consent received a blood sampling kit by mail. We obtained blood samples from 1,321 (73.3%) eligible breast cancer patients and 1,524 (74.1%) control women, but for technical reasons genotyping was not performed for 7 cases and 9 controls. Thus, genotype analysis was eventually performed for 1,314 cases and 1,515 controls.

Birth records

We attempted to identify birth records for breast cancer cases and controls, using the Swedish national registration numbers. Using information from the birth registry, research assistants visited 120 hospital and medical record archives to locate the original birth records for the study participants. Because this study included births occurring between 1918 and 1945 throughout all of Sweden, a large number of births at this time had occurred at home, and birth characteristic information had never been measured or recorded. We were able to locate birth records or midwife journals with adequate information on birth characteristics for about 50% of the participants.

Using a detailed form, perinatal characteristic information from available birth records was abstracted. Birth weight in grams and twin status were consistently listed in birth records. Less consistently, information was provided on birth length, head circumference and placental weight. In total, there were 710 breast cancer cases and 770 controls with both genotype information and birth weight data. We excluded 15 cases and 22 controls who were twin members and/or had missing information on one or more covariates. Eventually, there were 695 breast cancer cases and 748 control women with birth characteristic data and information on at least one of the evaluated SNPs.

Genotyping

We isolated DNA from 3 ml of whole blood with the Wizard Genomic DNA Purification Kit (Promega, Madison, WI) in accordance with the manufacturer's instructions. The current analysis focuses on the 7 SNPs identified from the BCAC study that replicated with genome-wide significance: rs2981582, rs12443621, rs8051542, rs889312, rs3817198, rs13281615, and rs3803662. Genotyping of these SNPs was conducted by Sequenom iPLEX and Taqman [8].

Statistical analysis

Initially, we examined the association between breast cancer risk and each of the following SNPs: rs2981582 (FGFR2), rs3803662 (TNRC9/LOC643714), rs889312 (MAP3K1), rs13281615 (8q24), rs3817198 (LSP1), rs12443621 (TNRC9/LOC643714), and rs8051542 (TNRC9/LOC643714). We considered these seven SNPs because they were identified in the BCAC genome-wide association study and were genome-wide significant after replication [8]. For each SNP, odds ratios (OR) and 95% confidence intervals (CI) for breast cancer risk were determined using logistic regression. To determine whether there was a linear trend with increasing number of variant alleles in each SNP, we calculated p values from Wald statistics including a continuous term in the model. Tests for interaction were based on a Wald test after inclusion of an interaction term. Data analysis was conducted using the SAS statistical software version 9.1 (SAS Institute, Cary, NC, USA). All p values presented are from two-sided tests of statistical significance.

Results

Table 1 shows the distribution of cases and controls by important breast cancer risk factors. As expected, cases were more likely to be nulliparous, have lower parity, have more previous benign breast disease, and more prior use of postmenopausal hormone therapy relative to controls. There was no difference in the age distributions because, as indicated, cases and controls were frequency matched on age. In these data, the established associations of breast cancer risk with age at menarche, body mass index (BMI), and birth weight were not evident.

Four of the seven breast cancer susceptibility loci identified from the BCAC genome-wide association study were significantly associated with breast cancer risk in this population with the strongest association for rs2981582, in line with previous reports [8, 9] (Table 2). Women homozygous for this allele had an almost twofold (OR = 1.99, 95% CI 1. 45 2.72) increased risk of breast cancer relative to women homozygous for the wild-type allele, whereas women heterozygous for this SNP were at slightly higher risk of breast cancer compared to homozygotes for the wild-type allele (OR = 1.25, 95% CI 0.99 1.57). Of the remaining six polymorphisms, trend tests with increasing number of high-risk alleles were statistically significant in three instances (rs12443621, rs13281615, and rs3803662), whereas in the three other instances, trends were in the direction expected on the basis of the BCAC study [8], but statistically not significant.

Because rs2981582 in intron 2 of FGFR2 has been repeatedly identified as the top hit in GWAS of breast cancer [8, 9] and has demonstrated the largest excess risk in our study, we first examined the interaction between rs2981582 and birth weight in relation to breast cancer

Table 1 Distribution of breast cancer cases and control women by selected breast cancer risk factors

	Cases (n 693)	Controls (n 747)	p value[a]
Age at study entry, years			0.76
<55	89 (12.8)	99 (13.3)	
55 <60	204 (29.4)	228 (30.5)	
60 <65	163 (23.5)	163 (21.8)	
65+	237 (34.2)	257 (34.4)	
BMI, kg m^{-2}			0.52
<20	35 (5.1)	38 (5.1)	
20 <25	325 (46.9)	349 (46.7)	
25 <30	241 (34.8)	281 (37.6)	
30+	92 (13.3)	79 (10.6)	
Parity			0.0003
Nulliparous	92 (13.3)	74 (9.9)	
1 Child	152 (21.9)	137 (18.3)	
2 Children	280 (40.4)	292 (39.1)	
3+ Children	169 (24.4)	244 (32.7)	
Previous benign breast disease			0.003
Yes	101 (14.6)	71 (9.5)	
No	592 (85.4)	676 (90.5)	
Postmenopausal hormone therapy use			0.01
Yes	267 (38.5)	240 (32.1)	
No	426 (61.5)	507 (67.9)	
Age at menarche, years[b]			0.58
<12	53 (8.3)	45 (6.6)	
12	112 (17.6)	118 (17.2)	
13	168 (26.4)	203 (29.5)	
14	175 (27. 5)	179 (26.0)	
15+	128 (20.1)	142 (20.7)	
Birth weight (g)			0.73
≤2,500	27 (3.9)	29 (3.9)	
2,501 ≤3,000	109 (15.7)	119 (15.9)	
3,001 ≤3,500	260 (37.5)	290 (38.8)	
3,501 ≤4,000	216 (31.2)	223 (29.9)	
>4,000	81 (11.7)	86 (11.5)	

[a] p value from chi square test with one degree of freedom

[b] Information on age at menarche was not available for 57 cases and 60 controls

risk (Table 3). To increase power without invoking linear trends, we grouped heterozygotes with homozygotes for the wild-type allele, because the difference in risk between homozygotes for the variant allele and heterozygotes is substantially larger than that between heterozygotes and homozygotes for the wild-type allele. We found a marginally significant interaction between birth weight and rs2981582 (p for interaction 0.07). As shown in Table 3, the interaction is generated by a strong, though marginally significant, positive association between birth weight and breast cancer risk among homozygotes for the variant allele (p for trend 0.07) and a weak, and clearly non-significant, inverse association between birth weight and breast cancer risk among heterozygotes and homozygotes for the wild-type allele (p for trend 0.68). Among women with the highest category of birth weight (>4,000 g), being homozygous for the rs2981582 variant was associated with an over fivefold increased risk of breast cancer (OR = 5.52, 95% CI 1.96 15.56) relative to women with wild-type alleles and similar birth weight.

We also examined a potential interaction between birth weight and the other 6 SNPs (Table 3), following the same approach used for rs2981582. Although none of the other SNPs demonstrated a significant interaction with birth weight, we did note a consistently positive association between birth weight and breast cancer risk among women homozygous for the high-risk allele.

Table 2 Odds ratios (OR) and 95% confidence intervals (CI) for the association between breast cancer susceptibility loci and breast cancer risk

	Cases[a]	Controls[a]	OR (95% CI)
rs2981582			
GG	240	319	1.0 (REF)
GA	304	324	1.25 (0.99 1.57)
AA	136	91	1.99 (1. 45 2.72)
p trend			<0.0001
rs12443621			
AA	193	241	1.0 (REF)
AG	337	366	1.15 (0.90 1.46)
GG	151	130	1.45 (1.07 1.96)
p trend			0.02
rs8051542			
GG	194	220	1.0 (REF)
GA	359	380	1.07 (0.84 1.36)
AA	132	135	1.11 (0.82 1.51)
p trend			0.49
rs889312			
TT	336	389	1.0 (REF)
TG	287	294	1.13 (0.91 1.41)
GG	57	54	1.22 (0.82 1.82)
p trend			0.19
rs3817198			
AA	320	365	1.0 (REF)
AG	300	314	1.09 (0.88 1.36)
GG	60	58	1.18 (0.80 1.74)
p trend			0.31
rs13281615			
AA	175	161	1.0 (REF)
AG	263	277	1.16 (0.91 1.48)
GG	223	273	1.33 (1.01 1.76)
p trend			0.04
rs3803662			
GG	333	415	1.0 (REF)
GA	300	273	1.37 (1.10 1.70)
AA	54	50	1.35 (0.89 2.03)
p trend			0.008

[a] Numbers do not add up because not all polymorphisms were determined in all study subjects

Discussion

In this population-based study, we evaluated the associations with breast cancer risk of seven SNPs previously identified in three GWAS [8 10]. In the current study, the association with breast cancer risk was statistically significant with respect to four of these SNPs, more strongly so as expected with the rs2981582 genotype, which has repeatedly been identified as the top hit in GWAS. After

Table 3 Odds ratios (OR) and 95% confidence intervals (95% CI) for breast cancer according to birth weight and susceptibility SNP genotype

	WT/WT and WT/VAR	VAR/VAR
rs2981582		
Birth weight (g)	OR(95% CI)	OR(95% CI)
	Cases/controls	Cases/controls
≤2,500	1.21 (0.64 2.29)	0.86 (0.30 2.53)
	21/20	6/8
2,501 ≤3,000	0.89 (0.63 1.26)	2.15 (1.12 4.13)
	80/103	28/15
3,001 ≤3,500	1.0 (REF)	1.27 (0.80 2.02)
	211/243	44/40
3,501 ≤4,000	1.02 (0.77 1.34)	1.90 (1.10 3.30)
	174/197	38/23
>4,000	0.84 (0.57 1.23)	4.61 (1.70 12.49)
	58/80	20/5
p value for trend	0.68 (negative trend)	0.07 (positive trend)
p value for interaction	0.07	
rs12443621		
Birth weight (g)		
≤2,500	1.36 (0.73 2.51)	0.87 (0.24 3.11)
	23/22	4/6
2,501 ≤3,000	1.24 (0.88 1.76)	0.92 (0.48 1.76)
	90/94	17/24
3,001 ≤3,500	1.0 (REF)	2.03 (1.34 3.08)
	185/240	72/46
3,501 ≤4,000	1.28 (0.96 1.70)	1.21 (0.76 1.92)
	171/174	42/45
>4,000	1.03 (0.70 1.51)	2.31 (1.00 5.34)
	61/77	16/9
p value for trend	0.78 (negative trend)	0.30 (positive trend)
p value for interaction	0.29	
rs8051542		
Birth weight (g)		
≤2,500	1.63 (0.86 3.09)	0.41 (0.11 1.53)
	24/18	3/9
2,501 ≤3,000	1.14 (0.81 1.60)	1.09 (0.55 2.16)
	92/99	17/19
3,001 ≤3,500	1.0 (REF)	1.66 (1.10 2.51)
	190/232	68/50
3,501 ≤4,000	1.26 (0.95 1.66)	0.80 (0.48 1.33)
	184/179	28/43
>4,000	1.07 (0.72 1.58)	1.40 (0.66 2.93)
	63/72	16/14
p value for trend	0.92 (negative trend)	0.81 (positive trend)

Table 3 continued

	WT/WT and WT/VAR	VAR/VAR
p value for interaction	0.80	
rs889312		
Birth weight (g)		
≤2,500	1.15 (0.65 2.03)	
	26/26	0/3
2,501 ≤3,000	1.09 (0.79 1.50)	0.69 (0.25 1.93)
	102/108	6/10
3,001 ≤3,500	1.0 (REF)	1.36 (0.75 2.46)
	229/263	26/22
3,501 ≤4,000	1.08 (0.83 1.41)	1.68 (0.81 3.47)
	194/206	19/13
>4,000	1.03 (0.72 1.49)	1.15 (0.37 3.61)
	72/80	6/6
p value for trend	0.91 (negative trend)	0.13 (positive trend)
p value for interaction	0.14	
rs3817198		
Birth weight (g)		
≤2,500	1.05 (0.58 1.89)	1.14 (0.28 4.59)
	23/25	4/4
2,501 ≤3,000	1.04 (0.76 1.44)	1.01 (0.38 2.66)
	100/109	8/9
3,001 ≤3,500	1.0 (REF)	1.08 (0.55 2.10)
	235/267	18/19
3,501 ≤4,000	1.15 (0.88 1.50)	0.89 (0.47 1.69)
	197/195	18/23
>4,000	0.89 (0.62 1.29)	4.55 (1.27 16.30)
	65/83	12/3
p value for trend	0.92 (negative trend)	0.26 (positive trend)
p value for interaction	0.27	
rs13281615		
Birth weight (g)		
≤2,500	1.50 (0.72 3.13)	0.55 (0.22 1.37)
	18/14	7/15
2,501 ≤3,000	1.26 (0.86 1.86)	0.72 (0.42 1.24)
	79/73	26/42
3,001 ≤3,500	1.0 (REF)	1.16 (0.81 1.65)
	143/167	103/104
3,501 ≤4,000	1.29 (0.93 1.78)	0.89 (0.60 1.32)
	143/130	63/83
>4,000	1.19 (0.77 1.84)	0.97 (0.54 1.74)
	55/54	24/29
p value for trend	0.97 (positive trend)	0.46 (positive trend)

Table 3 continued

	WT/WT and WT/VAR	VAR/VAR
p value for interaction	0.56	
rs3803662		
Birth weight (g)		
≤2,500	1.15 (0.64 2.06)	0.29 (0.03 2.60)
	25/25	1/4
2,501 ≤3,000	1.05 (0.76 1.45)	1.30 (0.49 3.42)
	100/110	9/8
3,001 ≤3,500	1.0 (REF)	1.65 (0.92 2.95)
	230/265	30/21
3,501 ≤4,000	1.13 (0.87 1.47)	0.89 (0.38 2.06)
	202/206	10/13
>4,000	1.07 (0.75 1.53)	1.15 (0.29 4.66)
	76/82	4/4
p value for trend	0.76 (positive trend)	0.87 (positive trend)
p value for interaction	0.94	

WT wild type; *VAR* high risk variant

grouping heterozygotes with homozygotes for the wild-type allele, we found a marginally significant interaction ($p \sim 0.07$) between birth weight and the rs2981582 genotype with respect to breast cancer risk. For the six remaining SNPs, there was an apparently stronger positive association between birth weight and breast cancer risk among women homozygous for the high-risk alleles than among carriers or homozygotes for the wild-type allele, although relatively large *p* values and multitude of comparisons hinder firm conclusions. These results suggest that the association of genetic susceptibility with breast cancer risk might be stronger among women with a larger mammary stem cell pool, as this is reflected in higher birth weight [19].

It has been suggested that the documented association of breast cancer risk with birth weight [12, 15, 16] could reflect the underlying association of this risk with the size of the mammary stem cell pool [19] and eventually mammary gland mass [17, 18]. In women who carry high-risk alleles, the association of birth weight (as a correlate of mammary stem cell pool) with breast cancer risk would be expected to be stronger. Our findings are compatible with this hypothesis, although the study was perhaps not sufficiently powered to document interactions of modest strength. Of note, in our study, associations were not statistically significant for the main effect of three of the seven SNPs identified in the context of GWAS [8]. When our study was initiated, the effect size of SNPs that could

be related to breast cancer was not known, precluding reliable power calculations.

The evidence that the association between birth weight and breast cancer could be modified by genetic susceptibility was stronger for the SNP in FGFR2 than for any of the other susceptibility loci. SNPs in intron 2 of FGFR2 have emerged as top hits from multiple GWAS of breast cancer [8, 9] and have been significantly associated with breast cancer risk in a number of populations including Europeans [8, 9], Asians [8, 23], Ashkenazi Jewish [24], and African American women [25]. FGFR2, a tyrosine kinase receptor belonging to a family of genes involved in growth and proliferation, is overexpressed in breast tumors [26] and may function as an oncogene [27, 28]. In the current study, we did observe a marginally significant positive association ($p \sim 0.07$) between birth weight and breast cancer among women who were homozygous for the risk allele of FGFR2. An interpretation of these results is that susceptibility loci in an oncogene such as FGFR2 put individual cells at higher risk of malignant transformation. To the extent that birth weight is a proxy for glandular mass, having an increased number of mammary stem cells with high-risk alleles would be associated with an increased risk of breast cancer.

Four additional loci not included in this study have recently been identified from further genotyping efforts in the BCAC [29] and CGEMS [30] genome-wide association studies and more are likely to be identified with pooling of GWASs. These additional loci have more modest effect sizes than FGFR2, and it is unlikely that loci will be discovered from subsequent pooling efforts with the same magnitude of association with breast cancer as has been demonstrated with FGFR2 [30]. With the possible exception of the FGFR2 gene, we know little about the mechanisms by which other susceptibility loci influence breast carcinogenesis. In this context, it would be of interest to examine whether carriers of the major breast cancer genes, BRCA1 and BRCA2, are at a disproportionally high risk for breast cancer if born with high birth weight.

A possible interaction between genetic susceptibility loci and birth weight in relation to breast cancer risk has not been previously investigated. In fact, there are few reliable studies of gene environment interactions and breast cancer risk. Our results are compatible with the hypothesis that the pool of mammary stem cells is critical in the intrauterine roots of breast cancer risk in adult life. This is because the size of the pool of mammary stem cells, as reflected in birth weight, appears to interact with genetic susceptibility in modulating breast cancer risk.

Acknowledgments This work was supported by the US Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs [Innovator Award W81XWH 05 1 0314].

References

1. Hankinson S, Tamimi R (2008) Hunter D. Breast cancer. In: Adami HO, Hunter D, Trichopoulos D (eds) Textbook of cancer epidemiology, 2nd edn. Oxford University Press, New York, pp 403 445
2. Lichtenstein P, Holm NV, Verkasalo PK et al (2000) Environ mental and heritable factors in the causation of cancer analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med 343:78 85
3. Lux MP, Fasching PA, Beckmann MW (2006) Hereditary breast and ovarian cancer: review and future perspectives. J Mol Med 84:16 28
4. Pharoah PD, Antoniou AC, Easton DF, Ponder BA (2008) Po lygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med 358:2796 2803
5. Pharoah PD, Antoniou A, Bobrow M et al (2002) Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 31:33 36
6. Ponder BA, Antoniou A, Dunning A et al (2005) Polygenic inherited predisposition to breast cancer. Cold Spring Harb Symp Quant Biol 70:35 41
7. Hirschhorn JN, Daly MJ (2005) Genome wide association studies for common diseases and complex traits. Nat Rev Genet 6:95 108
8. Easton DF, Pooley KA, Dunning AM et al (2007) Genome wide association study identifies novel breast cancer susceptibility loci. Nature 447:1087 1093
9. Hunter DJ, Kraft P, Jacobs KB et al (2007) A genome wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 39:870 874
10. Stacey SN, Manolescu A, Sulem P et al (2007) Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estro gen receptor positive breast cancer. Nat Genet 39:865 869
11. Trichopoulos D (1990) Hypothesis: does breast cancer originate in utero? Lancet 335:939 940
12. Michels KB, Xue F (2006) Role of birthweight in the etiology of breast cancer. Int J Cancer 119:2007 2025
13. Lagiou P (2007) Intrauterine factors and breast cancer risk. Lancet Oncol 8:1047 1048
14. Xue F, Michels KB (2007) Intrauterine factors and risk of breast cancer: a systematic review and meta analysis of current evi dence. Lancet Oncol 8:1088 1100
15. dos Santos Silva I, De Stavola B, McCormack V (2008) Col laborative group on pre natal risk factors and subsequent risk of breast cancer. Birth size and breast cancer risk: re analysis of individual participant data from 32 studies. PLoS Med 5:e193
16. Park SK, Kang D, McGlynn KA et al (2008) Intrauterine envi ronments and breast cancer risk: meta analysis and systematic review. Breast Cancer Res 10:R8
17. Adami H O, Persson I, Ekbom A et al (1995) The aetiology and pathogenesis of human breast cancer. Mutat Res 333:29 35
18. Trichopoulos D, Lagiou P, Adami HO (2005) Towards an inte grated model for breast cancer etiology: the crucial role of the number of mammary tissue specific stem cells. Breast Cancer Res 7:13 17
19. Strohsnitter WC, Savarese TM, Low HP et al (2008) Correlation of umbilical cord blood haematopoietic stem and progenitor cell levels with birth weight: implications for a prenatal influence on cancer risk. Br J Cancer 98:660 663

20. Magnusson C, Baron J, Persson I et al (1998) Body size in dif ferent periods of life and breast cancer risk in post menopausal women. Int J Cancer 76:29 34
21. Magnusson CM, Persson IR, Baron JA et al (1999) The role of reproductive factors and use of oral contraceptives in the aetiol ogy of breast cancer in women aged 50 to 74 years. Int J Cancer 80:231 236
22. Magnusson C, Baron JA, Correia N et al (1999) Breast cancer risk following long term oestrogen and oestrogen progestin replacement therapy. Int J Cancer 81:339 344
23. Liang J, Chen P, Hu Z et al (2008) Genetic variants in fibroblast growth factor receptor 2 (FGFR2) contribute to susceptibility of breast cancer in Chinese women. Carcinogenesis 29:2341 2346
24. Gold B, Kirchhoff T, Stefanov S et al (2008) Genome wide association study provides evidence for a breast cancer risk locus at 6q22.33. Proc Natl Acad Sci U S A 105:4340 4345
25. Udler MS, Meyer KB, Pooley KA et al (2009) FGFR2 variants and breast cancer risk: fine scale mapping using African American studies and analysis of chromatin conformation. Hum Mol Genet 18:1692 1703
26. Penault Llorca F, Bertucci F, Adélaïde J et al (1995) Expression of FGF and FGF receptor genes in human breast cancer. Int J Cancer 61:170 176
27. Theodorou V, Kimm MA, Boer M et al (2007) MMTV inser tional mutagenesis identifies genes, gene families and pathways involved in mammary cancer. Nat Genet 39:759 769
28. Katoh M (2008) Cancer genomics and genetics of FGFR2 (Review). Int J Oncol 33:233 237
29. Ahmed S, Thomas G, Ghoussaini M et al (2009) Newly dis covered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet 41:585 590
30. Thomas G, Jacobs KB, Kraft P et al (2009) A multistage genome wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). Nat Genet 41:579 584

British Journal of Cancer (2009) 100, 1794 1798

www.bjcancer.com

Insulin-like growth factor levels in cord blood, birth weight and breast cancer risk

P Lagiou[1,2,3], CC Hsieh[1,3,4], L Lipworth[5], E Samoli[2], W Okulicz[6], R Troisi[7], B Xu[8], P Hall[3], A Ekbom[9], HO Adami[1,3] and D Trichopoulos*[,1,3]

[1]Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA; [2]Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, University of Athens, 75M. Asias Street, Athens GR 115 27, Greece; [3]Department of Medical Epidemiology and Biostatistics, Box 281, Karolinska Institutet, Stockholm SE 171 77, Sweden; [4]Division of Biostatistics and Epidemiology, UMass Cancer Center, University of Massachusetts Medical School, 364 Plantation Street, LRB 427, Worcester, MA 01605, USA; [5]International Epidemiology Institute, 1455 Research Blvd, Suite 550, Rockville, MD 20850, USA; [6]Department of Physiology, ILAT Steroid RIA Laboratory, University of Massachusetts Medical School, Worcester, MA 01605, USA; [7]Division of Cancer Epidemiology and Genetics, Department of Health and Human Services, National Cancer Institute, National Institutes of Health,6120 Executive Boulevard, MSC 7242, Bethesda, MD 20892 7242, USA; [8]Department of Epidemiology, School of Public Health, Fudan University, Shanghai 200023, China; [9]Clinical Epidemiology Unit, Department of Medicine, Karolinska Institutet/Karolinska University Hospital, Stockholm SE 171 76, Sweden

Breast cancer incidence and birth weight are higher among Caucasian than Asian women, and birth size has been positively associated with breast cancer risk. Pregnancy hormone levels, however, have been generally lower in Caucasian than Asian women. We studied components of the insulin like growth factor (IGF) system in cord blood from 92 singleton babies born in Boston, USA, and 110 born in Shanghai, China, in 1994 1995. Cord blood IGF 1 was significantly higher among Caucasian compared with Chinese babies ($P < 10^{-6}$). The opposite was noted for IGF 2 ($P \sim 10^{-4}$). IGF 1 was significantly positively associated with birth weight and birth length in Boston, but not Shanghai. In contrast, stronger positive, though statistically non significant, associations of IGF 2 with birth size were only evident in Shanghai. The associations of birth weight and birth length were positive and significant in taller women (for IGF 1 in Boston $P \sim 0.003$ and 0.03, respectively; for IGF 2 in Shanghai $P \sim 0.05$ and ~ 0.04, respectively), among whom maternal anthropometry does not exercise strong constraints in foetal growth. The documentation of higher cord blood levels of IGF 1, a principal growth hormone that does not cross the placenta, among Caucasian than in Asian newborns is concordant with breast cancer incidence in these populations.
British Journal of Cancer (2009) **100,** 1794 1798. doi:10.1038/sj.bjc.6605074 www.bjcancer.com
Published online 5 May 2009

Keywords: insulin like growth factor; IGF 1; IGF 2; breast cancer; birth weight; birth length

During the past few decades, the possibility has been investigated that breast cancer can have roots in early life, including the intrauterine period. By the late 2000s, birth size had become a predictor of breast cancer risk (Michels and Xue, 2006; Park *et al*, 2008; dos Santos Silva *et al*, 2008). We have hypothesised that high levels of mammotropic hormones during pregnancy favour the generation of mammary tissue specific stem cells in the offspring and that the pool of these cells is an important predictor of risk (Trichopoulos *et al*, 2005). Because such stem cells are difficult to isolate, haematopoietic stem and progenitor cells have been used as markers. Concentration of these cells in cord blood is strongly positively associated with both cord blood levels of insulin like growth factor 1 (IGF 1) (Savarese *et al*, 2007) and birth weight (Strohsnitter *et al*, 2008), suggesting that IGF 1 may be an important factor in the intrauterine origin of breast cancer.

*Correspondence: Dr D Trichopoulos;
E mail: dtrichop@hsph.harvard.edu
Received 6 March 2009; revised 28 March 2009; accepted 5 April 2009; published online 5 May 2009

The incidence of breast cancer (Ferlay *et al*, 2004), as well as birth weight (Wen *et al*, 1995; Lagiou *et al*, 2003), are higher among Caucasian women in western countries than in Asian women in the east. We hypothesised, therefore, that cord blood levels of IGF 1, which does not cross the placenta (Holmes *et al*, 1999), are higher among Caucasian than among Asian neonates, and that cord blood IGF 1 levels are positively associated with birth size. We evaluated this hypothesis, and also examined the role of IGF 2, a main component of the IGF system in foetal life, by studying cord blood samples from babies born to women in Boston, USA and Shanghai, China.

MATERIALS AND METHODS

Study participants were adult pregnant women and their offspring, recruited from maternity clinics affiliated with two centres: Beth Israel Hospital in Boston, USA, and Shanghai Medical University in China. The study was approved by the Institutional Review Boards of the two centres, as well as the Institutional Review Boards of the Harvard School of Public Health and the US

Department of the Army. Details on the study have been reported earlier (Lipworth *et al*, 1999) and are summarised below.

At each centre, an authorised health professional met all pregnant women coming for their first routine prenatal visit to the collaborating maternity clinic, ascertained whether the woman was eligible to participate, explained to her the objectives of the study and the requirements for participation, and obtained informed consent. To be eligible, a pregnant woman had to be $<$ 40 years of age, have no more than one previous (liveborn or stillborn) child, be Caucasian in Boston and Chinese in Shanghai, and be able to understand and speak the local language. Women were excluded if they had taken any hormonal medication during the index pregnancy, if they had an earlier diagnosis of diabetes mellitus or thyroid disease, or if the foetus had a known major anomaly.

Between March 1994 and October 1995, among 402 women identified at Beth Israel Hospital in Boston, 304 pregnant women agreed to participate and met the eligibility criteria. In Shanghai, among 424 women identified between April 1994 and May 1995, 334 pregnant women agreed to participate and met the eligibility criteria. In both centres, gestational age was defined as the time since the first day of the last menstrual period. Baseline sociodemographic information was recorded and blood was drawn at gestational weeks 16 and 27. Cord blood was collected and additional information concerning the delivery and the newborn was ascertained from medical records and paediatric charts.

Cord blood was collected in sterile tubes without preservatives and refrigerated at 4°C for up to 24 h until centrifugation, after which, serum from each sampling was separated and distributed into aliquots. In Shanghai, cord blood samples were transported in a cooler to a laboratory near Shanghai Medical University. Serum aliquots were stored at 20°C for 5 7 days in the laboratory before being transported to Shanghai Medical University and stored at 80°C. All samples were shipped by air on dry ice to Boston where they were stored at 80°C together with the cord blood samples from Boston. We were able to collect cord blood samples of acceptable quality and sufficient quantities for the determinations of cord blood hormones for 202 uncomplicated full term pregnancies (37 42 weeks long without pregnancy toxaemia), 92 in Boston and 110 in Shanghai.

Hormone determinations were conducted at the ILAT Steroid RIA Laboratory of the University of Massachusetts Medical School. IGF 1, IGF 2, and IGFBP 3 were measured by coated tube immunoradiometric assay kits (Diagnostic System Laboratories, Inc., Webster, TX, USA). The laboratory estimated inter assay and intra assay coefficients of variation were, respectively, 9.0% and 3.3% for IGF 1, 5.9% and 3.4% for IGF 2, and 8.0% and 4.8% for IGFBP 3. There was no detectable cross reactivity of the IGF 1 assay with IGF 2 according to the manufacturer's specificity assessment.

Statistical analyses

Statistical analyses were conducted using the SPSS statistical package (Statistical Package for Social Sciences v. 16, Chicago, IL, USA). Multiple regression models were used to compare hormone levels between Boston and Shanghai controlling for maternal age, height, duration of gestation and weight gain (all continuously), as well as for parity and gender of offspring. The association of IGF 1 and IGF 2 with birth weight and birth length were examined by modelling the data through multiple regression models with, alternatively, birth weight and length as the outcomes, controlling also mutually for IGF 1 and IGF 2, as well as for IGFBP 3. Analyses were conducted separately for each centre, for all women, as well as for tall and short women, with the cut off for height set *a priori* at 163 cm (median height in Boston, third quartile in Shanghai) based on the results of an earlier study (Lagiou *et al*, 2005). The hypothesis is that, among shorter women, maternal anthropometry imposes constraints on birth size (Wen *et al*, 1995; Lagiou *et al*, 2003) and could trigger negative feedback mechanisms that could obscure positive associations of hormones with birth size.

RESULTS

Characteristics of mothers and offspring and cord blood levels of IGF 1, IGF 2, and IGFBP 3 are shown in Table 1. Caucasian mothers in Boston were significantly older and taller than Chinese mothers in Shanghai. In Boston, mothers were in almost equal proportions primi and biparae, whereas in Shanghai virtually all were primiparae. The number of boys and girls was identical in Boston, but there were substantially more boys than girls in Shanghai. Birth weight and length were higher in Boston. Table 1 also shows unadjusted mean values and standard deviations of the three studied components of the IGF axis. Cord blood levels of IGF 1 were higher in Boston than in Shanghai, whereas the opposite was evident with respect to IGF 2, while IGFBP 3 levels were higher in Shanghai.

In Table 2, cord blood levels of IGF 1, IGF 2, and IGFBP 3 are compared between cities adjusting for maternal age, height, weight gain, parity, duration of gestation, and gender of offspring. Levels of IGF 1 were significantly and substantially higher in Boston, whereas IGF 2 levels were significantly lower. The differences in IGF 1 and IGF 2 were amplified after adjustment both mutually as well as for IGFBP 3. As essentially all women in Shanghai were primiparae, we repeated the analyses for primiparae women only and the results were essentially unchanged.

Table 3 shows multiple regression derived partial regression coefficients of birth weight (upper panel) and birth length (lower panel) on one standard deviation increments of cord blood levels of IGF 1 and IGF 2 in Boston and Shanghai, both overall and in strata defined by maternal height. When all women within each centre were studied, results were adjusted for maternal age, height, parity, weight gain, duration of gestation, and gender of offspring. In the stratified analyses, in which the number of observations was

Table 1 Characteristics[a] of mothers and their offspring and cord blood levels of IGF 1, IGF 2, and IGFBP 3

	Boston (*n* 92)	**Shanghai (*n* 110)**
Age (years)	31.0 (3.0)	25.1 (3.3)
Maternal height (cm)	164.1 (7.2)	160.2 (4.9)
Maternal height		
⩽1.63 m	46 (50.0%)	81 (73.6%)
>1.63 m	46 (50.0%)	29 (26.4%)
Parity		
1	48 (52.2%)	109 (99.1%)
2	44 (47.8%)	1 (0.9%)
Duration of gestation (weeks)	40.1 (1.1)	40.0 (1.1)
Maternal weight gain (kg)[b]	11.5 (3.9)	8.9 (4.5)
Gender of offspring		
Male	46 (50.0%)	64 (58.2%)
Female	46 (50.0%)	46 (41.8%)
Birth weight (g)	3557.7 (490.2)	3492.6 (459.8)
Birth length (cm)	50.6 (2.5)	49.8 (3.1)
IGF 1 (ng ml ^{1})	98.4 (37.8)	79.0 (48.9)
IGF 2 (ng ml ^{1})	492.5 (100.3)	587.8 (140.1)
IGFBP 3 (ng ml ^{1})	2419.1 (1696.8)	3265.3 (2186.4)

IGF = insulin like growth factor. Uncomplicated full term singleton pregnancies in Boston, USA and Shanghai, China. [a]For continuous variables mean (s.d.); for categorical variables *n* (%). [b]Until the 27th week of gestation.

Epidemiology

Table 2 Percent differences[a] of cord blood levels of IGF 1, IGF 2, and IGFBP 3 between newborns in Boston, USA (reference) and Shanghai, China

	Unadjusted for the other IGF components		Adjusted for the other IGF components	
	Shanghai *vs* Boston (%)	***P*-value**	**Shanghai *vs* Boston (%)**	***P*-value**
IGF 1	39.6 (54.3, 20.2)	0.0005	50.9 (61.0, 38.1)	$<10^{-6}$
IGF 2	20.7 (7.7, 35.2)	0.001	23.2 (10.9, 37.0)	0.0001
IGFBP 3	23.0 (8.1, 64.7)	0.162	22.5 (5.1, 57.9)	0.118

IGF = insulin like growth factor. [a]Adjusted for maternal age, height and weight gain, parity, duration of gestation, and gender of offspring. Hormone levels were log transformed, so that the coefficients express percentage differences between centres.

Table 3 Multiple regression derived partial regression coefficients b[a] (and 95% confidence intervals, CIs) of birth weight (upper panel) and birth length (lower panel) on one standard deviation increments of cord blood levels of IGF 1, IGF 2, and IGFBP 3 in Boston, USA and Shanghai, China, overall and by maternal height. Statistically significant *P* values (<0.05) are indicated in bold fonts.

	All women				**Women ⩽1.63 m height**				**Women >1.63 m height**			
	Boston (*n* 92)		**Shanghai (*n* 110)**		**Boston (*n* 46)**		**Shanghai (*n* 81)**		**Boston (*n* 46)**		**Shanghai (*n* 29)**	
	***b* (CI)**	***P*-value**	***b* (CI)**	***P*-value**	***b* (CI)**	***P*-value**	***b* (CI)**	***P*-value**	***b* (CI)**	***P*-value**	***b* (CI)**	***P*-value**
Birth weight												
IGF 1 (per 44.4 ng ml^{-1})	141.7 (11.3, 272.1)	**0.03**	−3.7 (−133.5, 126.0)	0.96	−39.6 (−252.3, 173.1)	0.71	−1.2 (−139.9, 137.4)	0.99	260.4 (92.6, 428.1)	**0.003**	−33.6 (−420.6, 353.3)	0.86
IGF−2 (per 132.1 ng ml^{-1})	7.0 (−131.4, 145.3)	0.92	57.3 (−51.3, 166.0)	0.30	45.5 (−137.2, 228.1)	0.62	−2.3 (−124.1, 119.5)	0.97	1.0 (−196.7, 198.7)	0.99	292.8 (0.59, 584.9)	**0.05**
Birth length												
IGF 1 (per 44.4 ng ml^{-1})	0.85 (0.19, 1.51)	**0.01**	0.00 (−0.91, 0.91)	0.99	0.29 (−0.70, 1.28)	0.56	−0.21 (−1.39, 0.98)	0.73	1.14 (0.09, 2.19)	**0.03**	0.81 (−0.19, 1.81)	0.11
IGF 2 (per 132.1 ng ml^{-1})	0.10 (−0.60, 0.80)	0.77	0.65 (−0.11, 1.41)	0.09	0.44 (−0.41, 1.29)	0.30	0.73 (−0.31, 1.77)	0.17	−0.17 (−1.41, 1.06)	0.78	0.78 (0.02, 1.53)	**0.04**

IGF = insulin like growth factor; CI = confidence interval. [a]Adjusted for maternal age, height and weight gain, parity, duration of gestation, and gender of offspring in models for all women; adjusted for maternal age and weight gain, and duration of gestation in models by maternal height. In all models, IGF 1 and IGF 2 were adjusted both mutually and for IGFBP 3.

considerably reduced, parity and offspring gender, which mini mally affected within centre estimates, were not included among the covariates. In all models, IGF 1 and IGF 2 were adjusted both mutually and for IGFBP 3. In Boston, with respect to both birth weight and birth length, there were significant positive associations with IGF 1; stratified analyses indicated that these differences were generated exclusively by the offspring of taller women. In contrast, in Shanghai, IGF 1 was suggestively positively associated only with birth length and only among taller women ($P \sim 0.11$).

The results for IGF 2 were strikingly different. In Boston, no association was evident with respect to either birth weight or birth length, neither among women overall nor among taller or shorter women. In Shanghai, however, cord blood IGF 2 was suggestively positively associated with birth length ($P \sim 0.09$), whereas, among taller women, it was significantly positively associated with both birth weight and length. The cut off of 163 cm for maternal height was set *a priori*. Nevertheless, with respect to birth weight, the *P* value for interaction of maternal height with IGF 1 (both continuously) was 0.09 for Boston, whereas the *P* for interaction of maternal height with IGF 2 was 0.01 in Shanghai.

DISCUSSION

In our study involving pregnancies of 110 Asian women in China and 92 Caucasian women in USA, we have found that cord blood IGF 1 was significantly higher in Boston compared with Shanghai ($P < 10^{-6}$), whereas the opposite was noted with respect to IGF 2 ($P \sim 10^{-4}$) (Table 2). IGF 1 was positively associated with both birth weight and birth length among newborns in Boston, but not in Shanghai. With respect to IGF 2 in relation to birth size, there were suggestive positive associations in Shanghai mostly with respect to birth length ($P \sim 0.09$) (Table 3).

In an earlier study (Lagiou *et al*, 2005), we evaluated the association of pregnancy estriol in maternal sera with birth weight after stratification of women by stature. The results supported our hypothesis that, because among shorter women maternal anthro pometry imposes stronger constraints on birth size (Wen *et al*, 1995; Lagiou *et al*, 2003), negative feedback mechanisms might be triggered that masked positive associations of hormones with birth size. Similarly, in this study, the positive associations of cord blood IGF with birth weight and length were significant among taller women for IGF 1 in Boston and for IGF 2 in Shanghai, and essentially null among women of shorter stature.

Strengths of this investigation are the inclusion of participants from two populations with contrasting incidence of breast cancer, the implementation of a uniform protocol, the use of state of the art assays in a qualified laboratory, and the appreciable study size for a study of this nature. Limitations include lack of measurements of other IGF binding proteins and IGF receptors.

Birth size is positively associated with breast cancer risk several decades later (Michels and Xue, 2006; Park *et al*, 2008; dos Santos Silva *et al*, 2008) and birth weight is higher among Caucasian newborns in the United States compared with Asian newborns in China (Wen *et al*, 1995; Lagiou *et al*, 2003). The documentation of higher cord blood levels of IGF 1 among Caucasian compared with Asian newborns is concordant with the higher incidence of breast cancer in western compared with eastern Asian populations (Ferlay *et al*, 2004) and compatible with the role IGF appears to play in breast cancer, at least among premenopausal women (Renehan *et al*, 2004; Schernhammer *et al*, 2005). Of note, IGF 1 does not cross the placenta (Holmes *et al*, 1999) and this was also supported by our data, in which the correlation coefficients

between cord blood IGF 1 and maternal IGF 1 were very low. Our finding of higher cord blood levels of IGF 1 among Caucasian compared with Asian newborns is of particular importance because, in the same dataset, maternal pregnancy estradiol and estriol (Lipworth *et al*, 1999), as well as cord blood estriol, androstenedione, and testosterone (Troisi *et al*, 2008) have been reported to be significantly higher among Chinese than among Caucasian women. In a small subsample of 52 US and 22 Chinese newborns from the same dataset, in which IGF 2 was not measured, there was no significant difference between the two groups with respect to cord blood IGF 1 and, possibly by chance owing to the small subsample size, levels of IGF 1 appeared to be somewhat higher among Chinese newborns (Troisi *et al*, 2008).

Our results indicating that IGF 1 dominates foetal growth among Caucasians, whereas IGF 2 plays a similar role among Asians and that associations are evident among taller mothers are not directly comparable with previous studies, because mutual adjustment of IGF 1 and IGF 2 and stratification by maternal height were not generally undertaken. Nevertheless, cord blood IGF 1 has shown a consistently positive association with birth weight in Caucasian (Gluckman *et al*, 1983; Ashton *et al*, 1985; Ostlund *et al*, 1997; Ong *et al*, 2000; Christou *et al*, 2001) and less consistently in Asian populations (Wang *et al*, 1991; Yang and Kim, 2000; Yang and Yu, 2000; Lo *et al*, 2002; Hung *et al*, 2008). Cord blood IGF 2 associations with birth weight are generally weakly positive or null among both Caucasians (Gluckman *et al*, 1983; Ashton *et al*, 1985; Ong *et al*, 2000) and Asians (Lo *et al*, 2002; Pathmaperuma *et al*, 2007; Hung *et al*, 2008).

Cord blood IGF 1 has been shown to be positively associated with the size of the stem cell pool (Baik *et al*, 2005; Savarese *et al*, 2007), which has also been linked to birth size (Strohsnitter *et al*, 2008). The stem cell pool has been postulated to be related to breast cancer risk in later life (Trichopoulos *et al*, 2005). We found no reports concerning a possible association of this pool with IGF 2, which is a growth promoting hormone during gestation (O'Dell and Day, 1998). In the few analyses based on Asian populations, there was no relationship between birth size and breast cancer risk in later life (dos Santos Silva *et al*, 2008). The differential actions of IGF 1 and IGF 2 in embryonic life could be explained by the fact that both IGFs are known to bind to the signalling IGF 1 receptor, whereas IGF 2 also binds to the non signalling IGF 2 receptor (Ong *et al*, 2000).

Irrespectively of the underlying physiologic mechanisms, the difference in birth size between Caucasian and Asian newborn can account for only a small fraction of the differences in breast cancer incidence. However, the fact that endocrine perinatal influences on birth size are evident mostly, or exclusively, among newborn of taller women (Table 3) may explain the sharp contrast in breast cancer incidence between Caucasian and Asian women birth size is positively associated with adult height (Michels *et al*, 2006) and, over successive generations, improved nutrition, leading to increased adult body size, might reduce constraints on foetal growth and birth size (Wen *et al*, 1995; Lagiou *et al*, 2003), which in turn affects adult height. The cycle tends to repeat itself, notably over consecutive generations of Asians migrating to the west, who show a gradual increase of breast cancer incidence (Lagiou *et al*, 2003; Lagiou and Trichopoulos, 2008). Changes in age at first pregnancy, parity, and lactation also play a role in the increases of breast cancer among Asian migrants to western countries (Haenszel and Kurihara, 1968; Buell, 1973; Ziegler *et al*, 1993).

ACKNOWLEDGEMENTS

The study was funded by the W81XWH 05 1 0314 Innovator Award, US Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs.

REFERENCES

Ashton IK, Zapf J, Einschenk I, MacKenzie IZ (1985) Insulin like growth factors (IGF) 1 and 2 in human foetal plasma and relationship to gestational age and foetal size during midpregnancy. *Acta Endocrinol (Copenh)* **110:** 558 563

Baik I, Devito WJ, Ballen K, Becker PS, Okulicz W, Liu Q, Delpapa E, Lagiou P, Sturgeon S, Trichopoulos D, Quesenberry PJ, Hsieh CC (2005) Association of fetal hormone levels with stem cell potential: evidence for early life roots of human cancer. *Cancer Res* **65:** 358 363

Buell P (1973) Changing incidence of breast cancer in Japanese American women. *J Natl Cancer Inst* **51:** 1479 1483

Christou H, Connors JM, Ziotopoulou M, Hatzidakis V, Papathanassoglou E, Ringer SA, Mantzoros CS (2001) Cord blood leptin and insulin like growth factor levels are independent predictors of fetal growth. *J Clin Endocrinol Metab* **86:** 935 938

dos Santos Silva I, De Stavola B, McCormack V, Collaborative Group on Pre Natal Risk Factors and Subsequent Risk of Breast Cancer (2008) Birth size and breast cancer risk: re analysis of individual participant data from 32 studies. *PLoS Med* **5:** e193

Ferlay J, Bray F, Pisani P, Parkin DM (2004) GLOBOCAN 2002. Cancer Incidence, Mortality and Prevalence Worldwide. IARC CancerBase No. 5, version 2.0. IARC Press: Lyon

Gluckman PD, Johnson Barrett JJ, Butler JH, Edgar BW, Gunn TR (1983) Studies of insulin like growth factor I and II by specific radio ligand assays in umbilical cord blood. *Clin Endocrinol (Oxf)* **19:** 405 413

Haenszel W, Kurihara M (1968) Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. *J Natl Cancer Inst* **40:** 43 68

Holmes R, Porter H, Newcomb P, Holly JM, Soothill P (1999) An immunohistochemical study of type I insulin like growth factor receptors in the placentae of pregnancies with appropriately grown or growth restricted fetuses. *Placenta* **20:** 325 330

Hung TY, Lin CC, Hwang YS, Lin SJ, Chou YY, Tsai WH (2008) Relationship between umbilical cord blood insulin like growth factors and anthropometry in term newborns. *Acta Paediatr Taiwan* **49:** 19 23

Lagiou P, Hsieh CC, Trichopoulos D, Xu B, Wuu J, Mucci L, Tamimi R, Adami HO, Cnattingius S (2003) Birthweight differences between USA and China and their relevance to breast cancer aetiology. *Int J Epidemiol* **32:** 193 198

Lagiou P, Trichopoulos D (2008) Birth size and the pathogenesis of breast cancer. *PLoS Med* **5:** e194

Lagiou P, Samoli E, Lagiou A, Hsieh CC, Adami HO, Trichopoulos D (2005) Maternal height, pregnancy estriol and birth weight in reference to breast cancer risk in Boston and Shanghai. *Int J Cancer* **117:** 494 498

Lipworth L, Hsieh C c, Wide L, Ekbom A, Yu S z, Yu G p, Xu B, Hellerstein S, Carlstrom K, Trichopoulos D, Adami H O (1999) Maternal pregnancy hormone levels in an area with a high incidence (Boston, USA) and in an area with a low incidence (Shanghai, China) of breast cancer. *Br J Cancer* **79:** 7 12

Lo HC, Tsao LY, Hsu WY, Chen HN, Yu WK, Chi CY (2002) Relation of cord serum levels of growth hormone, insulin like growth factors, insulin like growth factor binding proteins, leptin, and interleukin 6 with birth weight, birth length, and head circumference in term and preterm neonates. *Nutrition* **18:** 604 608

Michels KB, Xue F (2006) Role of birthweight in the etiology of breast cancer. *Int J Cancer* **119:** 2007 2025

Michels KB, Xue F, Terry KL, Willett WC (2006) Longitudinal study of birthweight and the incidence of breast cancer in adulthood. *Carcinogenesis* **27:** 2464 2468

O'Dell SD, Day IN (1998) Insulin like growth factor II (IGF II). *Int J Biochem Cell Biol* **30:** 767 771

Ong K, Kratzsch J, Kiess W, Costello M, Scott C, Dunger D (2000) Size at birth and cord blood levels of insulin, insulin like growth factor I (IGF I), IGF II, IGF binding protein 1 (IGFBP 1), IGFBP 3, and the

soluble IGF II/mannose 6 phosphate receptor in term human infants. The ALSPAC Study Team. Avon Longitudinal Study of Pregnancy and Childhood. *J Clin Endocrinol Metab* **85:** 4266 4269

Ostlund E, Bang P, Hagenäs L, Fried G (1997) Insulin like growth factor I in fetal serum obtained by cordocentesis is correlated with intrauterine growth retardation. *Hum Reprod* **12:** 840 844

Park SK, Kang D, McGlynn KA, Garcia Closas M, Kim Y, Yoo KY, Brinton LA (2008) Intrauterine environments and breast cancer risk: meta analysis and systematic review. *Breast Cancer Res* **10:** R8

Pathmaperuma AN, Tennekoon KH, Senanayake L, Karunanayake EH (2007) Maternal and cord blood levels of insulin like growth factors I and II and insulin like growth factor binding protein 1: correlation with birth weight and maternal anthropometric indices. *Ceylon Med J* **52:** 48 52

Renehan AG, Zwahlen M, Minder C, O'Dwyer ST, Shalet SM, Egger M (2004) Insulin like growth factor (IGF) I, IGF binding protein 3, and cancer risk: systematic review and meta regression analysis. *Lancet* **363:** 1346 1353

Savarese TM, Strohsnitter WC, Low HP, Liu Q, Baik I, Okulicz W, Chelmow DP, Lagiou P, Quesenberry PJ, Noller KL, Hsieh CC (2007) Correlation of umbilical cord blood hormones and growth factors with stem cell potential: implications for the prenatal origin of breast cancer hypothesis. *Breast Cancer Res* **9:** R29

Schernhammer ES, Holly JM, Pollak MN, Hankinson SE (2005) Circulating levels of insulin like growth factors, their binding proteins, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* **14:** 699 704

Strohsnitter WC, Savarese TM, Low HP, Chelmow DP, Lagiou P, Lambe M, Edmiston K, Liu Q, Baik I, Noller KL, Adami HO, Trichopoulos D, Hsieh CC (2008) Correlation of umbilical cord blood haematopoietic stem and progenitor cell levels with birth weight: implications for a prenatal influence on cancer risk. *Br J Cancer* **98:** 660 663

Trichopoulos D, Lagiou P, Adami HO (2005) Towards an integrated model for breast cancer etiology: the crucial role of the number of mammary tissue specific stem cells. *Breast Cancer Res* **7:** 13 17

Troisi R, Lagiou P, Trichopoulos D, Xu B, Chie L, Stanczyk FZ, Potischman N, Adami HO, Hoover RN, Hsieh CC (2008) Cord serum estrogens, androgens, insulin like growth factor I, and insulin like growth factor binding protein 3 in Chinese and U.S. Caucasian neonates. *Cancer Epidemiol Biomarkers Prev* **17:** 224 231

Wang HS, Lim J, English J, Irvine L, Chard T (1991) The concentration of insulin like growth factor I and insulin like growth factor binding protein 1 in human umbilical cord serum at delivery: relation to fetal weight. *J Endocrinol* **129:** 459 464

Wen SW, Kramer MS, Usher RH (1995) Comparison of birth weight distributions between Chinese and Caucasian infants. *Am J Epidemiol* **141:** 1177 1187

Yang SW, Kim SY (2000) The relationship of the levels of leptin, insulin like growth factor I and insulin in cord blood with birth size, ponderal index, and gender difference. *J Pediatr Endocrinol Metab* **13:** 289 296

Yang SW, Yu JS (2000) Relationship of insulin like growth factor I, insulin like growth factor binding protein 3, insulin, growth hormone in cord blood and maternal factors with birth height and birthweight. *Pediatr Int* **42:** 31 36

Ziegler RG, Hoover RN, Pike MC, Hildesheim A, Nomura AM, West DW, Wu Williams AH, Kolonel LN, Horn Ross PL, Rosenthal JF, Hyer MB (1993) Migration patterns and breast cancer risk in Asian American women. *J Natl Cancer Inst* **85:** 1819 1827

Annals of Oncology Advance Access published October 13, 2010

original article

Annals of Oncology
doi:10.1093/annonc/mdq565

Maternal and cord blood hormone levels in the United States and China and the intrauterine origin of breast cancer

P. Lagiou[1,2], E. Samoli[2], W. Okulicz[3], B. Xu[4], A. Lagiou[5], L. Lipworth[6], C. Georgila[2], L. Vatten[7], H. O. Adami[1,8], D. Trichopoulos[1]* & C. C. Hsieh[1,9]

[1]Department of Epidemiology, Harvard School of Public Health, Boston, USA; [2]Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, University of Athens, Athens, Greece; [3]Department of Physiology, ILAT Steroid RIA Laboratory, University of Massachusetts Medical School, Worcester, USA; [4]Department of Epidemiology, School of Public Health, Fudan University, Shanghai, China; [5]Faculty of Health Professions, Athens Technological Educational Institute, Athens, Greece; [6]International Epidemiology Institute, Rockville, USA; [7]Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway; [8]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [9]Division of Biostatistics and Epidemiology, UMass Cancer Center, University of Massachusetts Medical School, Worcester, USA

Received 22 June 2010; accepted 16 August 2010

Background: Breast cancer is less common in China than in the United States and perinatal characteristics predict breast cancer risk in the offspring. We determined levels of pregnancy hormones in Boston and Shanghai to identify those possibly involved in the intrauterine origin of breast cancer.

Participants and methods: We compared maternal and cord blood levels of estradiol, estriol, testosterone, progesterone, prolactin, insulin-like growth factors (IGF) 1 and 2, insulin-like growth factor-binding protein 3, adiponectin and sex hormone-binding globulin (SHBG) in 241 Caucasian and 295 Chinese women.

Results: In both centers, hormone levels at the 16th were predictive of those at the 27th gestational week, but there was little correlation between maternal and cord blood levels. In cord blood, we found significantly ($P < 0.01$) higher levels of estradiol (44.2%), testosterone (54.5%), IGF-2 (22.7%) and strikingly SHBG (104.6%) in Shanghai women, whereas the opposite was true for IGF-1 (36.8%).

Conclusions: Taking into account the current understanding of the plausible biological role of the examined endocrine factors, those likely to be involved in the intrauterine origin of breast cancer are SHBG and IGF-2, with higher cord blood levels among Chinese, and IGF-1, with higher cord blood levels among Caucasian women.

Key words: breast cancer, cord blood, hormones, IGF, pregnancy, SHBG

introduction

The hypothesis that perinatal factors affect breast cancer risk was formally articulated in the early 1990s [1]. Since then, the accumulated evidence has linked two important perinatal factors with this risk, namely birth size [2 5] and pregnancy toxemia [3]. The underlying biological mechanisms are not known, but the endocrine environment during the perinatal period is thought to influence the risk of breast cancer in adulthood [6 10]. Our goal was to harvest information on early life endocrine factors which could account for the sharply higher incidence of breast cancer among Caucasian women in United States compared with Chinese women in China. To that end, we have determined levels of hormones with mammotropic potential in the maternal serum as well as the cord blood, in pregnancies of Caucasian women in Boston, USA and Chinese women in Shanghai, China. Although some results from this project have been previously reported [6, 11, 12], this paper presents new data on several additional hormones as well as an integrated picture of maternal and cord blood hormone levels in the two populations with contrasting breast cancer incidence.

materials and methods

subjects

From March 1994 to October 1995, all pregnant women coming for their first routine prenatal visit to the collaborating maternity clinics of the Beth Israel hospital in Boston, USA, and the Shanghai Medical University in China were met by authorized health professionals, who ascertained the woman's eligibility to participate, explained to her the objectives of the study and obtained informed consent [6]. A total of 402 Caucasian women in Boston, USA, and 424 Asian women in Shanghai, China, were identified. The study was approved by the Institutional Review Boards of the Beth Israel Hospital, Shanghai Medical University and Harvard School of Public

*Correspondence to: Prof. D. Trichopoulos, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. Tel: +1 617 432 4560; Fax: +1 617 566 7805; E mail: dtrichop@hsph.harvard.edu

Health. Eligibility criteria included age <40 years old, a maximum parity of two, absence of a prior diagnosis of diabetes mellitus or thyroid disease, no hormonal medication during the index pregnancy and no known fetal abnormality. The maximum parity of two criterion was imposed by the one child policy implemented in China and the need for comparability of the two cohorts.

Of the 402 eligible women in Boston, 77 refused to participate in one or more aspects of the study, 9 were subsequently excluded because of a spontaneous or induced abortion in the index pregnancy, 2 were excluded because of twin birth, whereas 10 were lost to follow up after the initial meeting. Of the remaining 304 women, we excluded 35 women with gestational age below 37 or above 42 weeks, 16 additional women with pregnancy toxemia and another 12 women with missing information. Eventually, 241 Caucasian women were considered in the present analysis. Additional exclusions were necessary due to limited availability of biological samples for some pregnancies.

Of the 424 eligible women in Shanghai, 15 refused to participate in one or more aspects of the study, 2 women were excluded owing to induced abortion in the index pregnancy and another two because of twin birth, whereas 7 women were lost to follow up leaving a total of 398 women. For 59 of those, no blood collection was accomplished. Of the remaining 339 women, we excluded 44 women who had gestation duration outside the range of 37 42 weeks inclusive. There were no Asian women with preeclampsia. Eventually, 295 Asian women were considered in the present analysis. Additional exclusions were necessary due to limited availability of biological samples for some pregnancies. It is noteworthy that, although exclusions were numerous, they were done for technical or administrative reasons, so their consequence is reduction of statistical power but not introduction of bias. Exclusions that could theoretically affect validity were those generated by losses to follow up and these were minimal (2.5% for women in Boston and 1.7% for women in Shanghai).

baseline data and laboratory measurements

Baseline sociodemographic and lifestyle information was recorded in interviews at the 16th and the 27th gestational week visit of the women to the clinic. Information about medical conditions was abstracted from the women's medical records. At delivery, additional information concerning the newborn, including duration of gestation and birth size parameters, was recorded. Detailed information concerning the study protocol and implementation has been published [6].

During the visits to the maternity clinics at the 16th and the 27th gestational week, 10 ml of venous blood was drawn from every woman at each visit. At delivery, cord blood was also collected. All blood samples were collected into sterile tubes without preservatives and stored at 4°C for up to 24 h until centrifugation. Samples were then centrifuged, and the serum was aliquoted and stored for hormonal assays at 80°C. In Shanghai, blood samples were transported in a cooler to a laboratory near Shanghai Medical University. Serum aliquots were stored at 20°C for 5 7 days in the laboratory before being transported to Shanghai Medical University and stored at 80°C. All samples were shipped by air on dry ice to Boston where they were stored at 80°C together with the samples from Boston.

Levels of estradiol (E2), estriol (E3), progesterone, sex hormone binding globulin (SHBG), testosterone, adiponectin, insulin like growth factor 1 (IGF 1) and insulin like growth factor binding protein 3 (IGFBP3) were measured in both the maternal sera and the cord blood. Prolactin was measured only in maternal sera and insulin like growth factor 2 (IGF 2) only in cord blood.

Measurements were conducted in two time periods. Maternal levels of E2, E3, SHBG, progesterone and prolactin were measured in the late 1990s at the Department of Clinical Chemistry of the Uppsala University Hospital in Sweden [6], whereas maternal levels of testosterone, adiponectin, IGF 1 and IGFBP3 as well as all hormone levels in cord blood were measured in 2006 at the ILAT Steroid RIA Laboratory of the University of Massachusetts Medical School. Measurements per each hormone were conducted simultaneously for samples from Boston and Shanghai.

Maternal estradiol 17b was measured with a time resolved competitive solid phase fluoroimmunoassay (AutoDELFIA Estradiol Kit; Wallac Oy, Turku, Finland), with laboratory imprecision 4.6% ± 0.8%. Maternal unconjugated E3 was measured with a similar time resolved competitive solid phase fluoroimmunoassay method (AutoDELFIA Unconjugated Oestriol Kit; Wallac Oy), with laboratory imprecision 8.0% ± 1.8%. Maternal SHBG was measured with a time resolved noncompetitive solid phase sandwich fluoroimmunoassay (AutoDELFIA SHBG Kit; Wallac Oy), with laboratory imprecision 4.8% ± 1.3%. Maternal progesterone was measured with a time resolved competitive solid phase fluoroimmunoassay (AutoDELFIA Progesterone Kit; Wallac Oy), with laboratory imprecision 1.7% ± 0.9%. Maternal prolactin was measured with a time resolved noncompetitive solid phase sandwich fluoroimmunoassay (AutoDELFIA Prolactin Kit; Wallac Oy), with laboratory imprecision 3.0% ± 0.3%. Maternal and cord blood testosterone was measured by radioimmunoassay kits from Diagnostic Products Corporation (DPC, Los Angeles, CA), with inter and intra assay coefficients of variation (CVs) of 8.9% and 5.6%, respectively. Maternal and cord blood adiponectin was measured by radioimmunoassay (Linco Research, St. Charles, MO) with inter and intra assay CV of 7.2% and 4.7%, respectively. Maternal and cord blood IGF 1 and IGFBP3 as well as cord blood IGF 2 were measured by coated tube immunoradiometric assay kits (Diagnostic System Laboratories, Inc., Webster, TX). The laboratory estimated inter and intra assay CVs were, respectively, 9.0% and 3.3% for IGF 1; 8.0% and 4.8% for IGFBP3 and 5.9% and 3.4% for IGF 2. There was no detectable cross reactivity of the IGF 1 assay with IGF 2 according to the manufacturer's specificity assessment. Cord blood E2 was measured by radioimmunoassay using kits from DPC (Los Angeles, CA). The inter assay CV was 6.8% and the intra assay CV was 3.4%. Cord blood unconjugated E3, progesterone and SHBG were measured using chemiluminescent immunoassay methodologies from DPC (Los Angeles, CA). The inter assay CVs were 9.2%, 7.9% and 4.8% and the intra assay CVs were 6.6%, 6.3% and 2.0%, respectively.

statistical analyses

The statistical analyses were conducted using the SPSS statistical package (Statistical Package for Social Sciences v. 16.0, Chicago, IL). At first, we distributed Caucasian and Chinese women by maternal offspring characteristics and we estimated the mean values, standard deviations, as well as percentiles of the studied endocrine compounds measured in maternal sera at the 16th and 27th gestational weeks and in the cord blood. We also calculated three Spearman correlation coefficients for each endocrine compound, between the 16th week maternal sera measurement, the 27th week maternal sera measurement and the cord blood measurement. Finally, we used multiple regression to compare log transformed cord blood and maternal serum hormone levels measured at the 16th and 27th gestational week between Boston (referent) and Shanghai, controlling for maternal age, maternal height, body mass index (BMI) before pregnancy and weight gain, as well as parity, duration of gestation, exact gestational week for maternal sampling and gender of offspring.

results

The characteristics of women and their singleton offspring in Boston, MA, USA, and Shanghai, China are shown in Table 1. In comparison to women in Boston, women in Shanghai were younger, of shorter stature, with lower prepregnancy BMI and generally primiparous (on account of the one child policy in China), whereas their offspring were more frequently boys and of lower birth weight. These factors were accounted for in comparisons of hormones between centers.

Tables 2 and 3 present central values and measures of dispersion for maternal hormones at around the 16th and the 27th gestational week, respectively, in Boston, USA, and Shanghai, China. In both centers, values of estradiol, estriol, prolactin, progesterone and IGF 1 tend to considerably increase with the progression of the pregnancy, whereas the rate of increase is lower for SHBG and IGFBP3 and even lower for testosterone. Adiponectin does not appear to increase between the 16th and the 27th gestational week and may in fact be declining.

Table 1. Characteristics of women[a] and their singleton offspring in Boston, MA, USA, and Shanghai, China

	Boston (n = 241)	Shanghai (n = 295)
Age (years)	31.0 (3.1)	25.3 (3.7)
Parity		
1	143 (59.3)	290 (98.3)
2	98 (40.7)	5 (1.7)
Duration of gestation (weeks)	40.0 (1.2)	40.0 (1.1)
Gestational week at first measurement	16.7 (1.1)	17.0 (2.0)
Gestational week at second measurement	27.1 (1.7)	26.9 (1.5)
Maternal weight gain until the second measurement (kg)	11.5 (4.7)	8.6 (4.4)
Maternal height (cm)	164.3 (6.6)	160.1 (4.8)
Prepregnancy BMI (kg/m^2)	22.0 (3.0)	20.0 (2.2)
Gender of offspring		
Male	128 (53.1)	175 (59.3)
Female	113 (46.9)	120 (40.7)
Birth weight (g)	3568.8 (478.7)	3421.6 (433.0)

For continuous variables, mean (SD); for categorical variables, n (%).
[a]Women with pregnancy duration between 37 and 42 weeks and no pregnancy toxemia.
BMI, body mass index; SD, standard deviation.

Table 4 shows central values and measures of dispersion for hormones in the umbilical cord blood in Boston, USA, and Shanghai, China. Comparisons of maternal hormone levels with hormone levels at the umbilical cord are meaningful only with respect to hormones measured at the same laboratory with the same method in maternal and cord blood (i.e. testosterone, adiponectin, IGF 1 and IGFBP3). In both Boston and Shanghai, in comparison to maternal blood levels, cord blood levels of testosterone and adiponectin were considerably higher, whereas those of IGF 1 and IGFBP3 were considerably lower. Notwithstanding the different methods used in maternal and cord blood, levels of SHBG in both centers appeared to be substantially lower in the umbilical cord compared with maternal blood.

Spearman correlation coefficients between the levels of the hormones measured in maternal blood at the 16th and 27th gestational week and the levels of these hormones in the cord blood in the two settings are shown in Table 5. Since correlation coefficients are dimensionless, it is acceptable to calculate these coefficients, even when the measurements are done at different laboratories with different methods, provided the rankings within laboratories are not very different. It is evident that the measured hormones track during pregnancy, so that levels at the 16th gestational week are predictive of levels at the 27th gestational week and vice versa. The correlations are strong with respect to SHBG, estradiol and testosterone; modest with respect to progesterone and adiponectin and weaker with respect to estriol, IGF 1 and IGFBP3. In contrast, there was little or no correlation between hormone levels in maternal blood and levels of these hormones in cord blood.

In Table 6, the levels of hormones measured in maternal blood at the 16th and 27th gestational week and in the cord blood are compared between Boston (reference) and Shanghai. Differences are evident between centers for most of the hormones in maternal blood or in the cord blood. In maternal sera, levels of estriol and prolactin (during both the 16th and 27th gestational week) as well as estradiol and IGFBP3 (only during the 16th gestational week), and testosterone (only during the 27th gestational week) were significantly higher in Shanghai than in Boston, whereas the opposite was true for levels of IGF 1 and adiponectin (during both the 16th and 27th gestational week) as well as progesterone (only during the 27th gestational week); no significant differences were evident between the two centers with respect to SHBG at either gestational week. In the cord blood, levels of estradiol, testosterone, IGF 2, IGFBP3 and most strikingly SHBG were significantly higher in Shanghai, whereas the opposite was true for levels of adiponectin and IGF 1.

discussion

We have studied singleton pregnancies and offspring of Caucasian women in Boston, USA, and Chinese women in

Table 2. Maternal serum levels of the indicated hormones at around the 16th gestational week in Boston, USA ($n = 232$[a]), and Shanghai, China ($n = 279$[a])

	Mean	SD	Percentiles				
			5%	25%	50%	75%	95%
Boston							
Estradiol (ng/ml)	3.8	1.8	1.9	2.6	3.5	4.6	7.1
Estriol (ng/ml)	1.1	0.5	0.5	0.8	1.0	1.3	2.1
SHBG (nmol/l)	362.9	89.6	219.0	298.1	360.6	423.0	524.9
Prolactin (μg/l)	43.9	25.1	15.7	27.4	38.4	54.3	86.8
Progesterone (ng/ml)	41.7	9.9	27.3	34.2	41.1	47.7	59.5
Testosterone (ng/ml)	0.6	0.4	0.2	0.4	0.5	0.8	1.2
Adiponectin (μg/ml)	17.6	8.1	8.0	12.3	15.7	21.4	34.4
IGF 1 (ng/ml)	192.7	83.2	70.8	126.0	194.0	247.0	331.4
IGFBP3 (ng/ml)	6481.1	2508.5	3633.2	4157.0	6208.0	8691.0	10415.2
Shanghai							
Estradiol (ng/ml)	5.5	2.5	2.4	3.7	4.9	6.8	11.0
Estriol (ng/ml)	1.8	1.0	0.6	1.1	1.5	2.1	3.7
SHBG (nmol/l)	427.0	91.7	272.6	371.3	422.6	487.2	586.9
Prolactin (μg/l)	62.5	33.1	22.1	40.3	57.9	78.4	125.6
Progesterone (ng/ml)	44.6	10.8	27.6	37.1	43.7	51.2	64.7
Testosterone (ng/ml)	0.6	0.3	0.2	0.4	0.6	0.7	1.3
Adiponectin (μg/ml)	15.9	7.1	7.4	11.0	14.7	18.7	29.5
IGF 1 (ng/ml)	139.3	72.0	23.0	95.0	131.0	177.0	259.0
IGFBP3 (ng/ml)	6833.3	2094.4	3783.0	4885.0	6974.0	8343.0	10233.0

[a]For certain hormones, slightly fewer number of samples (<10%) were available for analysis.
IGF 1, insulin like growth factor 1; IGFBP3, insulin like growth factor binding protein 3; SD, standard deviation; SHBG, sex hormone binding globulin.

Table 3. Maternal serum levels of the indicated hormones at around the 27th gestational week in Boston, USA ($n = 225$[a]), and Shanghai, China ($n = 281$[a])

	Mean	SD	Percentiles				
			5%	25%	50%	75%	95%
Boston							
Estradiol (ng/ml)	10.8	4.8	5.2	7.6	10.0	12.9	18.9
Estriol (ng/ml)	4.1	1.3	2.2	3.1	3.9	4.7	6.7
SHBG (nmol/l)	428.	111.8	258.2	348.9	420.2	500.1	646.4
Prolactin (μg/l)	92.4	36.7	42.1	65.8	89.9	111.7	147.5
Progesterone (ng/ml)	82.0	20.7	52.4	66.4	78.8	96.2	117.1
Testosterone (ng/ml)	0.7	0.3	0.3	0.4	0.6	0.8	1.2
Adiponectin (μg/ml)	15.8	7.3	7.4	10.8	14.0	19.8	28.3
IGF 1 (ng/ml)	238.4	110.1	87.3	160.0	231.5	302.8	442.2
IGFBP3 (ng/ml)	6818.2	2647.9	3678.7	4301.5	6878.0	8800.0	11275.8
Shanghai							
Estradiol (ng/ml)	13.1	5.0	7.6	9.9	12.2	15.3	22.3
Estriol (ng/ml)	6.4	2.8	3.4	4.7	5.9	7.3	10.3
SHBG (nmol/l)	475.4	123.7	297.5	389.4	454.3	543.8	714.7
Prolactin (μg/l)	115.9	38.0	61.7	90.0	112.1	138.0	185.8
Progesterone (ng/ml)	78.3	25.9	49.5	62.3	74.1	86.0	129.7
Testosterone (ng/ml)	0.8	0.4	0.4	0.5	0.7	0.9	1.4
Adiponectin (μg/ml)	12.0	5.1	5.9	8.5	10.9	14.2	22.1
IGF 1 (ng/ml)	206.1	85.9	87.2	140.0	194.0	260.0	371.6
IGFBP3 (ng/ml)	6385.9	2465.1	3425.6	4075.0	6182.0	8570.0	10229.3

[a]For certain hormones, slightly fewer number of samples (<12%) were available for analysis.
IGF 1, insulin like growth factor 1; IGFBP3, insulin like growth factor binding protein 3; SD, standard deviation; SHBG, sex hormone binding globulin.

Shanghai, China, and we determined a range of maternal blood hormones at the 16th and 27th gestational week, as well as in umbilical cord blood. We have established levels and the range of variation of the measured hormones in Caucasian and Chinese women and we have examined the possible concordance of the relative levels of these hormones with the sharp ecological contrast in the two populations with respect to breast cancer incidence. Results concerning maternal levels of

Table 4. Cord serum levels of the indicated hormones in Boston, USA (n = 92), and Shanghai, China (n = 110)

	Mean	SD	Percentiles				
			5%	25%	50%	75%	95%
Boston							
Estradiol (ng/ml)	35.4	18.7	10.9	22.4	30.0	44.3	82.1
Estriol (ng/ml)	382.9	160.2	173.6	273.5	360.5	462.0	738.1
SHBG (nmol/l)	42.0	42.1	19.3	26.0	31.6	38.6	126.1
Progesterone (ng/ml)	1078.5	771.7	384.6	574.8	878.0	1325.0	2565.1
Testosterone (ng/ml)	2.9	1.5	1.2	2.0	2.6	3.3	6.5
Adiponectin (μg/ml)	52.9	23.0	14.4	36.5	52.5	65.6	95.2
IGF 1 (ng/ml)	98.4	35.7	51.0	70.3	94.5	121.8	172.8
IGF 2 (ng/ml)	492.5	100.3	293.5	427.5	499.0	550.8	640.9
IGFBP3 (ng/ml)	2419.1	1696.8	437.7	1248.0	1992.0	3217.3	5948.8
Shanghai							
Estradiol (E2) (ng/ml)	79.7	75.5	21.7	44.2	61.3	87.0	251.3
Estriol (ng/ml)	406.9	135.7	168.0	318.0	395.0	490.0	628.0
SHBG (nmol/l)	116.1	117.1	24.7	33.1	51.7	175.3	374.9
Progesterone (ng/ml)	1058.3	685.4	176.4	652.7	875.5	1365.0	2818.0
Testosterone (ng/ml)	6.2	7.2	2.0	3.2	4.5	6.7	13.9
Adiponectin (μg/ml)	37.5	20.7	9.3	18.6	34.9	54.5	71.9
IGF 1 (ng/ml)	79.0	48.9	9.1	50.3	73.0	102.0	182.3
IGF 2 (ng/ml)	587.8	140.1	364.8	518.8	587.0	659.5	841.8
IGFBP3 (ng/ml)	3265.3	2186.4	1026.5	1809.8	2532.0	4112.8	7541.8

SHBG, sex hormone binding globulin; IGF, insulin like growth factor; IGFBP3, insulin like growth factor binding protein 3.

estrogens, SHBG, prolactin and progesterone have been previously published [6], as have results concerning umbilical cord blood levels of IGF 1, IGF 2 and IGFBP3 [12]. In contrast, hormone determinations conducted in the ILAT Steroid RIA Laboratory of the University of Massachusetts Medical School for maternal testosterone, adiponectin, IGF 1 and IGFBP3 as well as cord blood estradiol, estriol, SHBG, progesterone, testosterone and adiponectin have not been previously reported. Certain cord blood hormones were also measured in a different laboratory in smaller subsamples of the studied pregnancies [11]; no significant differences were noted with respect to cord blood levels of estradiol, estriol and IGF 1, whereas, in line with the findings of the present study, cord blood levels of testosterone and IGFBP3 were significantly higher among Chinese compared with Caucasian.

In addition to providing reliable data about the normal variations of the examined endocrine compounds in the two populations, this study allows examination of the correlations between the indicated hormones in the maternal blood at the 16th and 27th gestational week and in the cord blood in the two settings (Table 5). Levels of the hormones at the 16th gestational week were predictive of levels at the 27th gestational week and vice versa, whereas there was little or no correlation between hormone levels in maternal blood and levels of these hormones in cord blood.

We have attempted to interpret our findings in light of the considerably higher incidence of breast cancer in Caucasian compared with Chinese women, taking into account what is currently known about the role of these hormones in breast cancer etiology [13]. We have considered differences in cord blood levels as more relevant to the fetal endocrine environment and the possible long term breast cancer risk of the offspring.

In adult life, estrogens and testosterone have been consistently positively associated with breast cancer risk [13]. Both maternal and cord blood levels of estradiol, estriol and testosterone, however, were higher in Shanghai than in Boston by anywhere between 0.9% and 54.5% (Table 6). Thus, levels of these pregnancy hormones by themselves are unlikely to be critical actors in the intrauterine origin of breast cancer since, if anything, they are higher in the population with substantially lower breast cancer incidence. In contrast, levels of SHBG were sharply higher in the cord blood of women in Shanghai (by 104.64%), raising the possibility that these high levels in Chinese women may reduce the bioavailability of active endogenous estrogens and testosterone in the fetus. In adult life, SHBG in relation to breast cancer has been studied mainly as a modulator of the effects of estradiol and testosterone [14], but there is also evidence that high levels of this compound may be more directly associated with a reduction in breast cancer risk [15].

With respect to progesterone in adult life, results on its association with breast cancer risk have been inconsistent [13], with some studies suggesting an inverse association [16] and others indicating no association [17]. For prolactin, the evidence points to a positive association of adult life levels with breast cancer risk [18], whereas for adult life adiponectin, the reported results indicate an inverse association with this risk [19, 20]. In our study, maternal and cord blood progesterone and adiponectin levels were, if anything, lower, whereas maternal prolactin levels were

Table 5. Spearman correlation coefficients between maternal serum levels of the indicated hormones at the 16th and 27th gestational week and levels of these hormones in the cord blood in Boston, USA, and Shanghai, China

	Boston	Shanghai
Maternal serum measurement at weeks 16 and 27		
Estradiol	0.72**	0.61**
Estriol	0.37**	0.20*
SHBG	0.85**	0.67**
Progesterone	0.45**	0.39**
Testosterone	0.83**	0.62**
Adiponectin	0.46**	0.44**
IGF 1	0.25**	0.32**
IGFBP3	0.16*	0.32**
Maternal serum measurement at weeks 16 with cord blood measurement		
Estradiol	0.09	0.15
Estriol	−0.04	−0.06
SHBG	0.09	−0.09
Progesterone	−0.001	0.03
Testosterone	−0.001	0.05
Adiponectin	0.18	−0.12
IGF 1	0.02	−0.02
IGFBP3	0.19	0.01
Maternal serum measurement at weeks 27 with cord blood measurement		
Estradiol	0.17	0.12
Estriol	0.22*	0.00
SHBG	0.03	−0.12
Progesterone	0.09	0.06
Testosterone	0.16	0.17
Adiponectin	0.06	−0.20*
IGF 1	0.18	0.29*
IGFBP3	−0.02	−0.04

*$P < 0.05$; **$P < 0.001$.

SHBG, sex hormone binding globulin; IGF 1, insulin like growth factor 1; IGFBP3, insulin like growth factor binding protein 3.

higher, among Chinese compared with Caucasian women, sharply reducing the likelihood that any of these three hormones plays a major role in the intrauterine origin of breast cancer.

Adult life IGF 1 has been positively associated with breast cancer risk [21] and there is limited evidence that adult life IGF 2 may be inversely associated with this risk [22]. Because IGF 1 is substantially and significantly higher in maternal sera, and particularly in cord blood, among Caucasian women in the United States, whereas the opposite is true with respect to cord blood levels of IGF 2, it appears that these two components of the IGF system have distinct roles and that the balance of IGF 1 and IGF 2 in cord blood may be a key early life determinant of adult life breast cancer risk.

Strengths of our study are its longitudinal design, considerable sample size, measurement of a wide range of hormones and the implementation of an identical protocol in two population groups with sharply different breast cancer incidence rates, allowing meaningful ecological contrasts. A weakness of the study is that hormone determinations were done in two laboratories, but all comparisons between centers were for measurements undertaken at the same laboratory.

In conclusion, our study provides central values and measures of variation of a range of hormones during pregnancy in maternal and cord blood in Caucasian women in Boston and Chinese women in Shanghai. Taking into account the lower incidence of breast cancer among Chinese compared with Caucasian women and our current understanding of the role of the examined hormones in breast cancer risk, the endocrine factors likely to be involved in the intrauterine origin of breast cancer in the offspring are SHBG and IGF 2, with higher cord blood levels among Chinese, and IGF 1, with higher cord blood levels among Caucasian women.

Table 6. Comparison (%)[a] of the indicated hormones in maternal blood at the 16th and 27th gestational week and in the cord blood between Boston, USA (reference), and Shanghai, China

	16th gestational week[b]		27th gestational week[b]		Cord blood[c]	
	Shanghai versus Boston	*P*	Shanghai versus Boston	*P*	Shanghai versus Boston	*P*
Estradiol (ng/ml)	23.0 (10.9, 36.5)	<0.001	7.2 (−2.1, 17.3)	0.127	44.2 (12.5, 84.8)	0.004
Estriol (ng/ml)	34.9 (22.2, 49.0)	<0.001	42.9 (31.7, 55.0)	<0.001	2.6 (−13.2, 21.2)	0.775
SHBG (nmol/l)	4.2 (−1.7, 10.5)	0.170	−0.5 (−6.6, 6.0)	0.888	104.6 (47.2, 184.5)	<0.001
Prolactin (μg/l)	30.3 (14.9, 47.9)	<0.001	18.3 (8.1, 29.5)	<0.001		
Progesterone (ng/ml)	−3.2 (−8.7, 2.7)	0.274	−10.7 (−16.1, −4.8)	0.001	−20.4 (−40.1, 5.8)	0.120
Testosterone (ng/ml)	0.9 (−10.7, 14.0)	0.895	13.8 (1.3, 27.4)	0.023	54.5 (22.2, 95.4)	<0.001
Adiponectin (μg/ml)	−11.5 (−20.2, −1.8)	0.022	−23.8 (−31.3, −15.5)	<0.001	−26.9 (−43.0, −6.3)	0.015
IGF 1 (ng/ml)	−36.9 (−45.4, −27.10)	<0.001	−14.9 (−24.7, −3.9)	0.011	−36.8 (−52.2, −16.3)	0.002
IGF 2 (ng/ml)					22.7 (9.2, 38.0)	0.001
IGFBP3 (ng/ml)	12.2 (2.52, 22.80)	0.013	0.5 (−9.0, 11.0)	0.937	33.7 (0.1, 78.4)	0.049

[a]Hormone levels were log transformed; the coefficient expresses % difference between centers.

[b]Controlling for maternal age, BMI, height and weight gain, parity, duration of gestation, exact gestational week and gender of offspring.

[c]Controlling as in footnote b except for exact gestational week.

BMI, body mass index; IGF, insulin like growth factor; IGFBP3, insulin like growth factor binding protein 3; NA, not applicable; SD, standard deviation; SHBG, sex hormone binding globulin.

funding

W81XWH 05 1 0314 Innovator Award, US; Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs.

disclosure

The authors declare no conflict of interest.

references

1. Trichopoulos D. Hypothesis: does breast cancer originate in utero? Lancet 1990; 335: 939 940.
2. Michels KB, Xue F. Role of birthweight in the etiology of breast cancer. Int J Cancer 2006; 119: 2007 2025.
3. Xue F, Michels KB. Intrauterine factors and risk of breast cancer: a systematic review and meta analysis of current evidence. Lancet Oncol 2007; 8: 1088 1100.
4. dos Santos Silva I, De Stavola B, McCormack V; Collaborative Group on Pre Natal Risk Factors and Subsequent Risk of Breast Cancer. Birth size and breast cancer risk: re analysis of individual participant data from 32 studies. PLoS Med 2008; 5: e193.
5. Park SK, Kang D, McGlynn KA et al. Intrauterine environments and breast cancer risk: meta analysis and systematic review. Breast Cancer Res 2008; 10(1): R8.
6. Lipworth L, Hsieh C C, Wide L et al. Maternal pregnancy hormone levels in an area with a high incidence (Boston, USA) and in an area with a low incidence (Shanghai, China) of breast cancer. Br J Cancer 1999; 79: 7 12.
7. Wuu J, Hellerstein S, Lipworth L et al. Correlates of pregnancy oestrogen, progesterone and sex hormone binding globulin in the USA and China. Eur J Cancer Prev 2002; 11: 283 293.
8. Schernhammer ES. In utero exposures and breast cancer risk: joint effect of estrogens and insulin like growth factor? Cancer Causes Control 2002; 13: 505 508.
9. Troisi R, Potischman N, Johnson CN et al. Estrogen and androgen concentrations are not lower in the umbilical cord serum of pre eclamptic pregnancies. Cancer Epidemiol Biomarkers Prev 2003; 12: 1268 1270.
10. Xu B, Lipworth L, Wide L et al. Maternal and gestational correlates of pregnancy prolactin and growth hormone in USA and China. Eur J Cancer Prev 2003; 12: 35 42.
11. Troisi R, Lagiou P, Trichopoulos D et al. Cord serum estrogens, androgens, insulin like growth factor I, and insulin like growth factor binding protein 3 in Chinese and U.S. Caucasian neonates. Cancer Epidemiol Biomarkers Prev 2008; 17: 224 231.
12. Lagiou P, Hsieh CC, Lipworth L et al. Insulin like growth factor levels in cord blood, birth weight and breast cancer risk. Br J Cancer 2009; 100: 1794 1798.
13. Hankinson S, Tamimi R, Hunter D. Breast cancer. In Adami HO, Hunter D, Trichopoulos D (eds), Textbook of Cancer Epidemiology, 2nd edition. New York: Oxford University Press 2008; 403 445.
14. Fortunati N, Catalano MG, Boccuzzi G, Frairia R. Sex Hormone Binding Globulin (SHBG), estradiol and breast cancer. Mol Cell Endocrinol 2010; 316: 86 92.
15. Lipworth L, Adami HO, Trichopoulos D et al. Serum steroid hormone levels, sex hormone binding globulin, and body mass index in the etiology of postmenopausal breast cancer. Epidemiology 1996; 7: 96 100.
16. Micheli A, Muti P, Secreto G et al. Endogenous sex hormones and subsequent breast cancer in premenopausal women. Int J Cancer 2004; 112: 312 318.
17. Eliassen AH, Missmer SA, Tworoger SS et al. Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women. J Natl Cancer Inst 2006; 98: 1406 1415.
18. Tworoger SS, Eliassen AH, Rosner B et al. Plasma prolactin concentrations and risk of postmenopausal breast cancer. Cancer Res 2004; 64: 6814 6819.
19. Miyoshi Y, Funahashi T, Kihara S et al. Association of serum adiponectin levels with breast cancer risk. Clin Cancer Res 2003; 9: 5699 5704.
20. Mantzoros C, Petridou E, Dessypris N et al. Adiponectin and breast cancer risk. J Clin Endocrinol Metab 2004; 89: 1102 1107.
21. The Endogenous Hormones and Breast Cancer Collaborative Group. Insulin like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. Lancet Oncol 2010; 11(6): 530 542.
22. Walker K, Fletcher O, Johnson N et al. Premenopausal mammographic density in relation to cyclic variations in endogenous sex hormone levels, prolactin, and insulin like growth factors. Cancer Res 2009; 69: 6490 6499.

British Journal of Cancer (2008) 98, 660 663

www.bjcancer.com

Correlation of umbilical cord blood haematopoietic stem and progenitor cell levels with birth weight: implications for a prenatal influence on cancer risk

WC Strohsnitter[1], TM Savarese[2,*], HP Low[3], DP Chelmow[1], P Lagiou[4], M Lambe[5], K Edmiston[6], Q Liu[2], I Baik[2], KL Noller[1], H-O Adami[5,7], D Trichopoulos[7] and C-C Hsieh[*,2,5,7]

[1]Department of Obstetrics and Gynecology, Tufts New England Medical Center, Boston, MA, USA; [2]Department of Cancer Biology and Cancer Center, University of Massachusetts Medical School, Worcester, MA, USA; [3]Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA; [4]Department of Hygiene and Epidemiology, School of Medicine, Athens University, Athens, Greece; [5]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [6]Division of Hematology/Oncology, Department of Medicine, UMass Medical School and UMass Memorial Health Care, Worcester, MA, USA; [7]Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA

We examined the relation with birth weight and umbilical cord blood concentrations of haematopoietic stem and progenitor populations in 288 singleton infants. Across the whole range of birth weight, there was a positive relation between birth weight and $CD34^+CD38^-$ cells, with each 500 g increase in birth weight being associated with a 15.5% higher (95% confidence interval: 1.6 31.3%) cell concentration. $CD34^+$ and $CD34^+$c kit^+ cells had J shaped relations and CFU GM cells had a U shaped relation with birth weight. Among newborns with $\geqslant$3000 g birth weights, concentrations of these cells increased with birth weight, while those below 3000 g had higher stem cell concentrations than the reference category of 3000 3499 g. Adjustment for cord blood plasma insulin like growth factor I levels weakened the stem and progenitor cell birth weight associations. The positive associations between birth weight and stem cell measurements for term newborns with a normal to high birth weight support the stem cell burden hypothesis of cancer risk.
British Journal of Cancer (2008) **98,** 660 663. doi:10.1038/sj.bjc.6604183 www.bjcancer.com
Published online 5 February 2008

Keywords: birth weight; cancer risk; prenatal exposure; stem cell

The *in utero* environment and perinatal factors may influence cancer risk of the offspring later in life (Trichopoulos, 1990). One parameter that reflects *in utero*/perinatal influences, that is birth weight, has been positively correlated with subsequent risk of childhood cancer (Schüz and Forman, 2007) and, in adults, breast (Michels and Xue, 2006), prostate (Eriksson *et al*, 2007) and colorectal cancers (Nilsen *et al*, 2005) and, indeed, overall cancer risk (Ahlgren *et al*, 2007). Mechanistically, a 'stem cell burden' theory (Adami *et al*, 1995) has been proposed to account for the positive relationship between birth weight and the risk of certain cancers, especially that of the breast. By this hypothesis, the levels of *in utero*/perinatal mitogens and other factors determine the size of the stem cell pools in the developing fetus; elevated tissue stem cell numbers drive the formation of larger organs and hence might be associated with larger birth weights. The greater the stem cell pool size, however, the greater the chance that one of the stem cells will be mutated by a carcinogen, or undergo a DNA replicative error, initiating oncogenic transformation. Hence, individuals with high birth weights might be at greater lifetime cancer risk (Trichopoulos *et al*, 2005).

The stem cell burden theory predicts (1) that the *in utero* levels of particular mitogens should correlate positively with stem cell population levels and (2) that the stem cell levels should correlate positively with birth weight. In a previous study, we demonstrated that the umbilical cord blood concentrations of various haemato poietic stem and progenitor populations correlated with cord blood plasma levels of particular mitogens, especially insulin like growth factor 1 (IGF 1) (Savarese *et al*, 2007). Here, we determine whether or not these measurable haematopoietic stem cell/ progenitor values, serving as surrogates of overall stem cell potential, are positively associated with birth weight.

MATERIALS AND METHODS

The umbilical cord blood study protocol was approved by the institutional review boards of the American Red Cross, the University of Massachusetts Medical School, the University of Massachusetts/Memorial Health Care System, St. Vincent's

*Correspondence: Dr C C Hsieh, Division of Biostatistics and Epide miology, Department of Cancer Biology, UMass Cancer Center, University of Massachusetts Medical School, Worcester, MA, USA;
E mail: chung.hsieh@umassmed.edu
* Deceased.
Received 29 October 2007; revised 30 November 2007; accepted 7 December 2007; published online 5 February 2008

Hospital and Tufts New England Medical Center (T NEMC). Consenting study subjects were recruited from one of two sources: (1) participants in the Worcester, MA based American Red Cross cord blood program (ACBP), in which a haematopoietic stem cells from umbilical cord blood were collected for possible transplanta tion, from August 2002 to June 2003, and (2) pregnant women delivering at T NEMC from October 2004 to April 2006. All the cord blood samples were from full term (gestational age $\geqslant$37 weeks) singleton infants. The processing of samples, which includes the determination of cord blood volumes, the determina tion of initial levels of total nucleated cells (TNC) and mono nuclear cells (MNC) before centrifugations or manipulation, the quantitation of haematopoietic stem/progenitor cell populations and the determination of cord blood plasma hormone levels, have been described (Savarese *et al*, 2007). The haematopoietic stem and progenitor populations that were quantitated (1) $CD34^+$ cells, a heterogeneous population of early multipotent stem and progenitor cells, committed progenitors and differentiating cells (Xiao and Dooley, 2000); (2) $CD34^+CD38^-$ cells, which represent more primitive stem cells depleted of lineage committed pre cursors (Xiao and Dooley, 2000); (3) $CD34^+$c kit^+ cells, which also represent a more primitive stem cell population that has relatively high cloning efficiencies in semisolid culture (Sharkey *et al*, 1994; Mayani and Lansdorp, 1998); and (4) granulocyte macrophage colony forming units (CFU GM), a functional measure of the number of proliferative granulocyte/macrophage committed haematopoietic precursor cells (Abboud *et al*, 1992; Hoffbrand *et al*, 2001).

Birth weight was first studied as a categorical variable ($<$3000, 3000 3499, 3500 3999 and $\geqslant$4000 g). Geometric means for the stem cell measurements were estimated within the indicated categories of birth weight. Multivariate linear regression was used to examine the association between natural log transformed measures of stem cell potential (dependent variable) and birth weight (independent variable, using 3000 3499 g as the reference), adjusting for maternal and neonatal characteristics (mother's age, race of parents, number of live births, gestation duration, baby's gender, delivery time and study site). To assess whether there was an underlying linear trend, birth weight was next analyzed as a continuous variable across the whole range of birth weight values with the effect estimates expressed for each 500 g increase in birth weight. The fitted coefficients from the regression analyses were exponentiated to obtain the estimated proportional change in birth weight associated with each independent variable. Statistical significance was set at 0.05 (two sided). Levels of IGF 1, which had the strongest association with levels of stem cells among the hormones and growth factors examined in a previous analysis (Savarese *et al*, 2007), were further adjusted to explore its influence on the association between birth weight and stem cell measurements.

RESULTS

The characteristics of the study subjects are shown in Table 1. Subjects from the ACBP and T NEMC patient groups had similar age and gestation duration. Parental ethnicity was more varied in the T NEMC samples, while the ACBP subjects had higher parities, more male newborns and lower birth weights.

The associations were analysed in multivariate analysis adjust ing for maternal age, parental race, parity, gestation duration, gender, delivery time and study site (Table 2, upper panel). There was a J shaped association between birth weight categories and concentrations of TNC (lymphocytes, monocytes and granulo cytes), as well as a J shaped relation with MNC (lymphocytes and monocytes), both including more differentiated cells. Among the stem cell populations, there was a positive association with $CD34^+CD38^-$ cells across the whole range of birth weight categories, with each 500 g increase being associated with 15.5% higher levels of this cell population (95% confidence interval: 1.6, 31.3%). A J shaped relation was observed for the $CD34^+$ and $CD34^+$c kit^+ cells: for birth weights of 3000 g or greater, stem cell concentrations increased with birth weight, while the lowest category of $<$3000 g had higher levels than the category of 3000 3499 g. For CFU GM, an approximate U shaped relation was observed, with the lowest birth weight category having the highest levels of this cell population (Table 2, upper panel).

Adjusting for cord blood plasma levels of IGF 1 in the multivariate analysis weakens the association (Table 2, lower panel). The association with $CD34^+CD38^-$ remained positive but was no longer statistically significant: each 500 g increase in birth weight was associated with a 7.9% increase in this cell sub population (95% confidence interval: 6.2, 24.0%). The lowest weight category continued to have the highest CFU GM cells after adjusting for IGF 1 levels.

We conducted further analyses adjusting for other hormones and in samples from different ethnic groups. Adjusting for estriol or insulin like growth factor binding protein 3, which had statistically significant but weaker associations with cord blood levels of stem cells (Savarese *et al*, 2007), in place of IGF 1, had much less effect on the association with birth weight (data not shown). A linear relation between stem cell measurements and birth weight was observed for newborns whose parents were Caucasian, but did not have a consistent shape among samples in the mixed non Caucasian group (Table 3).

Table 1 Characteristics of the study subjects by study site

Characteristics	ACBP (*n* 39) Mean ± s.d. or *N* (%)	T-NEMC (*n* 249) Mean ± s.d. or *N* (%)
Mother's age (years)	29.7 ± 5.1	30.1 ± 5.6
Parity		
1	14 (35.9)	104 (45.0)
2	12 (30.8)	64 (27.7)
3	7 (18.0)	39 (16.9)
4 or more	6 (15.4)	24 (10.4)
Race/ethnicity of mother and biological father		
Both Caucasian	36 (92.3)	111 (52.4)
Both African American	1 (2.6)	19 (9.0)
Both Asian	0 (0.0)	39 (18.4)
Both Hispanic	0 (0.0)	14 (6.6)
Mixed	2 (5.1)	29 (13.7)
Gestation duration (weeks)	39.6 ± 1.4	39.6 ± 1.2
Newborn gender		
Male	21 (55.3)	113 (50.0)
Female	17 (44.7)	113 (50.0)
Birth weight (g)	3313.1 ± 450.6	3416.4 ± 428.2
$<$3000	11 (28.2)	40 (16.1)
3000 3499	15 (38.5)	110 (44.2)
3500 3999	11 (28.2)	77 (30.9)
$\geqslant$4000	2 (5.1)	22 (8.8)

Abbreviations: ACBP = American Red Cross Cord Blood Program; T NEMC = Tufts New England Medical Center. Data on some variables were unavailable for subjects with missing values.

DISCUSSION

The stem cell burden hypothesis has been invoked as an explanation for the positive link between birth weight and risk for both childhood and adult cancers (Adami *et al*, 1995). This hypothesis proposes that *in utero* environments that promote

Epidemiology

Table 2 Multiple linear regression analysis for the association between measurements of haematopoietic stem cell populations and birth weight

		Birth weight (g) in categories				Birth weight per 500 g
Analytic model	**Cell measurements**	**<3000 % Difference (95% CI)**	**3000–3499 Reference (geometric mean[a])**	**3500–3999 % Difference (95% CI)**	**⩾4000 % Difference (95% CI)**	**% Difference (95% CI)**
Adjusted for core covariates[b]	TNC	2.1 (8.5, 14.0)	0.0 (15.00)	7.8 (1.7, 18.3)	8.4 (8.1, 27.8)	3.0 (2.5, 8.9)
	MNC	5.7 (6.4, 19.3)	0.0 (6.91)	6.5 (3.9, 18.0)	12.5 (6.2, 34.9)	2.1 (3.9, 8.4)
	$CD34^+$[c]	2.4 (19.2, 29.7)	0.0 (7.04)	12.5 (8.1, 37.7)	20.8 (15.5, 72.8)	9.5 (2.6, 23.2)
	$CD34^+CD38^-$[c,d]	1.1 (23.6, 28.1)	0.0 (3.11)	19.1 (4.3, 48.1)	**47.9 (0.4, 117.8)**	**15.5 (1.6, 31.3)**
	$CD34^+$c *kit*$^+$	13.7 (14.0, 50.4)	0.0 (5.80)	18.9 (5.5, 49.6)	34.2 (10.4, 101.1)	10.0 (4.0, 26.1)
	CFU GM[e]	**37.0 (2.1, 83.8)**	0.0 (4.04)	29.0 (0.1, 66.6)	31.4 (16.2, 106.1)	6.1 (8.4, 23.0)
Adjusted for core covariates and IGF 1	TNC	0.6 (10.0, 12.5)	0.0 (15.00)	9.7 (0.4, 20.8)	12.3 (5.5, 33.4)	4.5 (1.5, 11.0)
	MNC	4.0 (8.0, 17.7)	0.0 (6.91)	8.5 (2.5, 20.6)	16.8 (3.5, 41.4)	3.7 (2.9, 10.8)
	$CD34^+$[c]	11.0 (12.3, 40.6)	0.0 (7.04)	2.5 (16.5, 25.8)	0.3 (31.0, 44.0)	0.6 (11.4, 14.2)
	$CD34^+CD38^-$[c,d]	5.2 (19.0, 36.6)	0.0 (3.11)	10.9 (11.5, 38.8)	27.6 (14.7, 90.9)	7.9 (6.2, 24.0)
	$CD34^+$c *kit*$^+$	22.8 (6.9, 62.1)	0.0 (5.80)	7.0 (15.3, 35.2)	10.5 (26.9, 66.9)	0.7 (12.9, 16.4)
	CFU GM[e]	**45.6 (8.6, 95.1)**	0.0 (4.04)	15.5 (11.6, 50.9)	11.5 (29.8, 77.1)	1.7 (16.2, 15.3)

[a]Unadjusted geometric means. TNC, initial total nucleated cells × 10^6 per ml; MNC, initial mononuclear cells × 10^6 per ml; the unit for the stem cell populations ($CD34^+$, $CD34^+CD38^-$, $CD34^+$c *kit*$^+$, and CFU GM) was per 1000 MNC. [b]Core covariates included mother's age, race of parents (both Caucasian or not), parity, gestation duration, gender of baby (male or female), delivery time (night or day) and study site (ACBP or T NEMC). [c]$n = 233$ with complete information on all the covariates. [d]Determined only in the T NEMC derived samples. [e]Data from the T NEMC derived samples on which this assay was conducted. Statistical significance of $P < 0.05$ are given in bold.

Table 3 Ethnic specific, core covariate adjusted[a] multiple linear regression analysis for the association between measurements of haematopoietic stem cell populations and birth weight

		Birth weight (g) in categories				Birth weight per 500 g
Ethnicity	**Stem cell measurements**	**<3000 % Difference (95% CI)**	**3000–3499 Reference (geometric mean[b])**	**3500–3999 % Difference (95% CI)**	**⩾4000 % Difference (95% CI)**	**% Difference (95% CI)**
Both parents Caucasian	$CD34^+$	13.9 (37.2, 18.0)	0.0 (7.73)	8.3 (15.2, 38.3)	37.5 (11.2, 113.1)	**16.6 (0.9, 34.7)**
	$CD34^+CD38^-$	18.0 (43.0, 17.9)	0.0 (3.41)	14.3 (13.7, 51.4)	**80.5 (9.1, 198.7)**	**23.8 (4.7, 46.4)**
	$CD34^+$c *kit*$^+$	11.5 (40.8, 32.3)	0.0 (6.44)	18.4 (11.0, 57.5)	63.4 (1.0, 169.5)	**22.9 (3.6, 45.8)**
	CFU GM	19.3 (25.9, 92.8)	0.0 (4.01)	36.9 (3.4, 94.1)	40.0 (25.3, 162.5)	19.7 (3.3, 48.2)
Either parent non Caucasian	$CD34^+$	28.4(12.1, 87.6)	0.0 (6.47)	19.9 (16.6, 72.3)	0.6 (46.6, 85.3)	0.5 (18.1, 23.5)
	$CD34^+CD38^-$	23.2 (16.2, 81.2)	0.0 (2.84)	28.2 (10.8, 84.4)	13.5 (39.1, 111.7)	6.3 (13.8, 31.0)
	$CD34^+$c *kit*$^+$	42.1 (6.2, 115.5)	0.0 (5.39)	22.2 (17.2, 80.4)	9.2 (43.9, 112.6)	0.4 (20.4, 24.7)
	CFU GM	**68.4 (14.0, 148.6)**	0.0 (4.08)	21.0 (18.1, 78.6)	24.3 (37.5, 147.1)	7.3 (26.0, 16.1)

[a]Core covariates included mother's age, parity, gestation duration, gender of baby (male or female), delivery time (night or day) and study site (ACBP or T NEMC). [b]Unadjusted geometric means. The unit for the stem cell measurements was per 1000 MNC. Statistical significance of $P < 0.05$ are given in bold.

expansion of stem cell pools result in infants with high birth weights; the larger the stem cell pool, the greater the risk that one of these stem cells will undergo malignant transformation. In support of the first tenet of this hypothesis, we have demonstrated that the concentrations of stem and progenitor cell populations in umbilical cord blood, serving as surrogates for overall stem cell potential, correlate with cord blood plasma levels of certain mitogens, notably IGF 1 (Savarese *et al*, 2007). The hypothesis also predicts that newborns with high birth weights should have elevated stem cell populations. Our findings indicate that there is a positive association between birth weight and haematopoietic stem cell measurements in the cord blood samples among newborns with normal to high (⩾3000 g) birth weights. This association is strongest with $CD34^+CD38^-$ cells, a relatively primitive haematopoietic stem cell population. These data are in line with previous studies, which showed a positive relationship between cord blood $CD34^+$ or CFU GM levels and birth weight (Shlebak *et al*, 1998; Ballen *et al*, 2001; Aroviita *et al*, 2004).

However, in our study, newborns in the lowest birth weight category (<3000 g) can have higher levels of stem/progenitor cell measurements than those with 3000–3499 g birth weight resulting in a J- or U-shaped relation between stem cell levels and birth weight. This finding is intriguing, as J- or U-shaped relationships have often been observed in childhood cancers (Schüz and Forman, 2007) and neurological (Schüz *et al*, 2001; Von Behren and Reynolds, 2003), prostate (Eriksson *et al*, 2007), colorectal (Nilsen *et al*, 2005) and early onset breast cancers (Sanderson *et al*, 1996; Innes *et al*, 2000; Mellemkjaer *et al*, 2003).

The possible elevated levels of stem cells at low birth weight warrant further investigations. Using birth weight to define small for gestation for full term healthy infants, a majority (>86%) of such infants undergo accelerated growth during the first 6–12 months after birth and attain normal height later in life (Karlberg and Albertsson Wikland, 1995). Premature infants (i.e., those born before the 33rd gestational week, generally with low birth weights) have been shown to have elevated foetal and cord blood $CD34^+$ and $CD34^+CD38^-$ levels relative to full term infants (most with a normal birth weight) (Shields and Andrews, 1998; Wyrsch *et al*, 1999). Relevant to this phenomenon, it has been reported that premature female newborns have an increased risk for breast cancer later in life (Ekbom *et al*, 2000). It is not known whether premature or small newborns harbour elevated stem/progenitor

Epidemiology

cell populations, because these populations have not had enough developmental time to undergo a normal course of differentiation, or if such infants build up a stem cell reserve for growth compensation during the postnatal period. In any case, our J shaped relation observed between birth weight and stem cell measurements requires further confirmation as ethnicity specific results showed a linear relation in the considerably larger group of Caucasian cord blood samples.

Finally, we examined cord blood plasma IGF 1 levels in relation to stem cell levels and birth weight. Insulin like growth factor 1 had a positive association with stem cell measurements in our samples (Savarese *et al*, 2007) and was strongly associated with birth weight (geometric means were 41.18, 57.86, 78.45 and 102.77 ng ml^{-1}, respectively, for the four categories of birth weight in Table 2 and in the literature (Bennett *et al*, 1983; Fant *et al*, 1993; Reece *et al*, 1994)). The growth hormone/IGF 1 axis has been suggested to serve as a master developmental regulator, coordinat ing stem cells in multiple organs (Ginestier and Wicha, 2007). When we controlled for cord plasma IGF 1 levels, the associations were considerably weakened. This would be expected if IGF 1 regulates stem cell potential and this is, in turn, a determinant of birth weight. To determine whether birth weight is an accurate reflection of stem cell potential, the focus should be on the results without adjusting for IGF 1.

We conclude that there is a J or U shaped positive association between birth weight and stem cell measurements, linear among term newborns with normal to high birth weight and stronger with the more primitive stem cell sub population. The nonlinear relationship, however, suggests that birth weight is not an unequivocal indicator of stem cell potential in the context of a prenatal origin of cancer risk. Quantitation of stem cell pools might prove to be a more accurate predictor of cancer risk than birth weight *per se*, in accordance with the stem cell burden hypothesis.

ACKNOWLEDGEMENTS

We are grateful for the effort of the attending physicians and residents of the Department of Obstetrics and Gynecology and Family Practice and the staff on the Labor and Delivery ward at the Tufts New England Medical Center. This study was support by a Grant (R01CA90902) from the US National Institutes of Health. PL, HOA, DT and CCH were supported in part by the W81XWH 05 1 0314 Innovator Award, USA. Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs.

REFERENCES

Abboud M, Xu F, LaVia M, Laver J (1992) Study of early hematopoietic precursors in human cord blood. *Exp Hematol* **20:** 1043 1047

Adami H O, Persson I, Ekbom A, Wolk A, Ponten J, Trichopoulos D (1995) The aetiology and pathogenesis of human breast cancer. *Mutat Res* **333:** 29 35

Ahlgren M, Wohlfahrt J, Olsen LW, Sørensen TIA, Melbye M (2007) Birth weight and risk of cancer. *Cancer* **110:** 412 419

Aroviita P, Teramo K, Hiilesmaa V, Westman P, Kekomaki R (2004) Birthweight of full term infants is associated with cord blood CD34+ cell concentration. *Acta Paediatr* **93:** 1323 1329

Ballen KK, Wilson M, Wuu J, Ceredona AM, Hsieh C, Stewart FM, Popovsky MA, Quesenberry PJ (2001) Bigger is better: maternal and neonatal predictors of hematopoietic potential of umbilical cord blood units. *Bone Marrow Transplant* **27:** 7 14

Bennett A, Wilson DM, Liu F, Nagashima R, Rosenfeld RG, Hintz RL (1983) Levels of insulin like growth factors I and II in human cord blood. *J Clin Endocrinol Metab* **57:** 609 612

Ekbom A, Erlandsson G, Hsieh CC, Trichopoulos D, Adami H O, Cnattingius S (2000) Risk of breast cancer in prematurely born women. *J Natl Cancer Inst* **92:** 840 841

Eriksson M, Wedel H, Wallander M A, Krakau I, Hugosson J, Carlsson S, Svärdsudd K (2007) The impact of birth weight on prostate cancer incidence and mortality in a population based study of men born in 1913 and followed up from 50 to 85 years of age. *The Prostate* **67:** 1247 1254

Fant M, Salafia C, Baxter RC, Schwander J, Vogel C, Pezzullo J, Moya F (1993) Circulating levels of IGFs and IGF binding proteins in human cord serum: relationships to intrauterine growth. *Regul Pept* **48:** 29 39

Ginestier C, Wicha MS (2007) Mammary stem cell number as a determinate of breast cancer risk. *Breast Cancer Res* **9:** 109

Hoffbrand AV, Pettit JE, Moss P (2001) *Essential Haematology* 4th edn, pp 1 11. Malden, MA: Blackwell Science

Innes K, Byers T, Schymura M (2000) Birth characteristics and subsequent risk for breast cancer in very young women. *Am J Epidemiol* **152:** 1121 1128

Karlberg J, Albertsson Wikland K (1995) Growth in full term small for gestational age infants: from birth to final height. *Pediatr Res* **38:** 733 739

Mayani H, Lansdorp PM (1998) Biology of human umbilical cord blood derived hematopoietic stem/progenitor cells. *Stem Cells* **16:** 153 165

Mellemkjaer L, Olsen ML, Sorensen HT, Thulstrup AAM, Olsen J, Olson JH (2003) Birth weight and risk of early onset breast cancer (Denmark). *Cancer Causes Control* **14:** 61 64

Michels KB, Xue F (2006) Role of birth weight in the etiology of breast cancer. *Int J Cancer* **119:** 2007 2025

Nilsen TIL, Romundstad PR, Troisi R, Potischman N, Vatten LJ (2005) Birth size and colorectal cancer risk: a prospective population based study. *Gut* **54:** 1728 1732

Reece EA, Wiznitzer A, Le E, Homko CJ, Behrman H, Spencer EM (1994) The relation between human fetal growth and fetal blood levels of insulin like growth factors I and II, their binding proteins and receptors. *Obstet Gynecol* **84:** 88 95

Sanderson M, Williams MA, Malone KE, Stanford JL, Emanuel I, White E, Daling JR (1996) Perinatal factors and risk of breast cancer. *Epidemiology* **7:** 34 37

Savarese TM, Strohsnitter WC, Low HP, Liu Q, Baik I, Okulicz W, Chelmow DP, Lagiou P, Quesenberry PJ, Noller KL, Hsieh CC (2007) Correlation of umbilical cord blood hormones and growth factors with stem cell potential: implications for the prenatal origin of breast cancer hypothesis. *Breast Cancer Res* **9:** R29

Schüz J, Forman MR (2007) Birthweight by gestational age and childhood cancer. *Cancer Causes Control* **18:** 655 663

Schüz J, Kaletsch U, Meinert R, Kaatsch P, Spix C, Michaelis J (2001) Risk factors for neuroblastoma at different stages of disease Results from a population based case control study in Germany. *J Clin Epidemiol* **54:** 702 709

Sharkey AM, Jokhi PP, King A, Loke YW, Brown KD, Smith SK (1994) Expression of c kit and kit ligand at the human maternofetal interface. *Cytokine* **6:** 195 205

Shields LE, Andrews RG (1998) Gestational age changes in circulating CD34+ hematopoietic stem/progenitor cells in fetal cord blood. *Am J Obstet Gynecol* **178:** 931 937

Shlebak AA, Roberts IAG, Stevens TA, Syzdlo RM, Goldman JM, Gordon MY (1998) The impact of antenatal and perinatal variables on cord blood haemopoietic stem/progenitor cell yield available for transplantation. *Br J Haematol* **103:** 1167 1171

Trichopoulos D (1990) Does breast cancer originate *in utero*? *Lancet* **335:** 939 940

Trichopoulos D, Lagiou P, Adami HO (2005) Towards an integrated model for breast cancer etiology: the crucial role of the number of mammary tissue specific stem cells. *Breast Cancer Res* **7:** 13 17

Von Behren J, Reynolds P (2003) Birth characteristics and brain cancers in young children. *Int J Epidemiol* **32:** 248 256

Wyrsch A, dale Carbonare V, Jansen W, Chklovskaia E, Nissen C, Surbek D, Holzgreve W, Tichelli A, Wodnar Filipowicz A (1999) Umbilical cord blood from preterm human fetuses is rich in committed and primitive hematopoietic progenitors with high proliferative and self renewal capacity. *Exp Hematol* **27:** 1338 1345

Xiao M, Dooley DC (2000) Cellular and molecular aspects of human $CD34^{+}CD38^{-}$ precursors: analysis of a primitive hematopoietic popula tion. *Leuk Lymphoma* **38:** 489 497

Accepted for publication by Annals of Oncology

Novel measurements of mammary stem cells in human umbilical cord blood as prospective predictors of breast cancer susceptibility in later life

L. Qiu[1§], H. P. Low[1§], C.-I. Chang[1], W. C. Strohsnitter[2], M. Anderson[3], K. Edmiston[3], H.-O. Adami[4,5], A. Ekbom[6], P. Hall[5], P. Lagiou[4,7], D. Trichopoulos[4], C.-C. Hsieh[1*]

[1]*Department of Cancer Biology, University of Massachusetts Medical School, Worcester, USA;*

[2]*Department of Obstetrics and Gynecology, Tufts Medical Center, Boston, USA;*

[3]*Division of Hematology/Oncology, Department of Medicine, University of Massachusetts Medical School and University of Massachusetts Memorial Health Care, Worcester, USA;*

[4]*Department of Epidemiology, Harvard School of Public Health, Boston, USA;*

[5]*Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden;*

[6]*Unit of Clinical Epidemiology, Department of Medicine, Karolinska Institute/Karolinska University Hospital, Stockholm, Sweden;*

[7]*Department of Hygiene, Epidemiology and Medical Statistics, University of Athens Medical School, Greece*

[§]L. Qiu and H. P. Low contributed equally to this work.

Correspondence to: Dr. C.-C. Hsieh, Department of Cancer Biology, Lazare Research Building, 4th Floor, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605, United States of America.

Tel: +1-508-856-4780; Fax: +1-508-856-1310; E-mail: chung-cheng.hsieh@umassmed.edu

Background: The size of the breast stem cell pool could underlie the intrauterine roots of breast cancer. We studied whether breast stem cells exist in umbilical cord blood and if they correlate with hematopoietic stem cell measurements that have been positively associated with perinatal risk factors for breast cancer.

Subjects and methods: We isolated mononuclear cells from umbilical cord blood of 170 singleton, full-term pregnancies, and determined, by reverse transcription polymerase chain reaction, the presence of genes of putative breast epithelial stem cell/progenitor markers (including EpCAM, CD49f (α6-integrin), CD117 (c-kit), CD24, and CD29 (β1-integrin)). By immunocytochemistry, we co-localized protein expressions of $EpCAM^+CD49f^+$, $CD49f^+CD24^+$ and $CD24^+CD29^+$. We correlated concentrations of putative breast stem cell/progenitor subpopulations, quantified by flow cytometry, with concentrations of hematopoietic stem cells.

Results: Mammary stem cell phenotypes were identified in umbilical cord blood. The measured $EpCAM^+$ subpopulation was positively correlated with concentrations of $CD34^+$ and $CD34^+CD38^-$ hematopoietic stem cells (both $P = 0.006$). Additionally, $EpCAM^+CD49f^+$ and $CD49f^+CD24^+$ subpopulations were positively correlated to the $CD34^+$ cells ($P = 0.03$ and 0.008, respectively).

Conclusion: The positive association between measurable breast and hematopoietic stem cells in human umbilical cord blood suggests plausible mechanisms for a prenatal influence on breast cancer risk.

Keywords: epithelial cell adhesion molecule, flow cytometry, hematopoietic stem cell, integrins, *in utero* environment, prenatal origin

INTRODUCTION

The hypothesis that a woman's risk for breast cancer in the adult life is influenced already in the *in utero* environment [1] has implicated a role of mitogens and stem cells [2, 3]. In a large cohort of normal, singleton pregnancies, we showed that the concentration of hematopoietic stem cells in umbilical cord blood is positively correlated with perinatal levels of insulin-like growth factor 1 (IGF-1) and estrogens [4, 5], and with birth weight [6], an indicator of adult life breast cancer risk [3, 7, 8]. These results are consistent with a "stem cell burden and susceptibility" hypothesis which predicts that levels of mitogens increase the number of stem cells ("burden"), and that such stem cells are targets for genetic and/or epigenetic alterations ("susceptibility") that might lead to malignant transformation [3, 9, 10]. In previous studies [4-6], we used hematopoietic stem cells (defined by the $CD34^+$, $CD34^+CD38^-$, and $CD34^+CD117^+$ cell surface markers) as a surrogate for the overall stem cell levels. However, levels of epithelial breast stem or progenitor cells in the *in utero* environment would be a more biologically relevant indicator for future breast cancer risk.

Although umbilical cord blood contains hematopoietic [11] and endothelial [12, 13] stem/progenitor cells, it is challenging to assume that organ-specific breast stem cells are present in an *in utero* compartment far removed from the organ of interest. However, umbilical cord blood mononuclear cells (MNC) express embryonic [14] and neural [15, 16] stem cell markers. Additionally, umbilical cord blood stem cells, possibly including mesenchymal stem cells, can be differentiated to other cell types, such as osteoblasts, chondroblasts, and adipocytes, indicating their pluripotent potential [17,

18]. Mammary phenotypes have however not been reported in human umbilical cord blood.

Here we analyzed MNC derived from human umbilical cord blood for gene expressions of markers reported for putative breast epithelial stem cells and progenitors [19] and quantified such cell populations by flow cytometry. We report the existence of putative breast stem/progenitor cell phenotypes in umbilical cord blood and, more importantly, that the concentrations of certain breast stem/progenitor cell phenotypes found in umbilical cord blood correlate positively with those of hematopoietic stem cells.

SUBJECTS AND METHODS

The study protocol was approved by the institutional review boards of the University of Massachusetts Medical School, Worcester, MA, USA, and Tufts Medical Center, Boston, MA, USA.

Subject recruitment and umbilical cord blood processing

Study subjects were recruited from November, 2006 to November, 2010 among pregnant women who delivered at the Tufts Medical Center, who were 18 years or older, HIV and hepatitis B-negative, with a fetus free of anomalies by ultrasound examination. For this analysis, we included only full term (gestational age ≥ 37 weeks), normotensive, and singleton pregnancies. Infants were delivered according to standard obstetrical practices. Umbilical cord blood was collected from the umbilical vein into a sterile bag containing 35 ml of citrate-phosphate-dextrose anticoagulant (Fenwal, Lake

Zurich, IL, USA). Samples were processed for MNC using a Ficoll-Paque density gradient within 24 hours of birth as described previously [5, 6].

Reverse transcription-polymerase chain reaction (RT-PCR)

PolyA$^+$ mRNA was isolated from MNC using the QuickPrep Micro mRNA Purification Kit (Amersham Biosciences, Piscataway, NJ, USA) and reversed transcribed to cDNA using the Superscript III First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA), according to manufacturer's instructions. Genes of interest were detected using a set of two specific forward and reverse primers for each gene in a polymerase chain reaction (PCR): EpCAM, forward 5'-TTGGTGATGAAGGCAGAAATGAATGG-3' and reverse 5'-TGAACTAAAACACAAAGCAAGAGAAAAACCT-3' giving a PCR product of 268 bp; CD49f [20], forward 5'-CAAGATGGCTACCCAGATAT-3' and reverse 5'-CTGAATCTGAGAGGGAACCA-3' giving a PCR product of 210 bp; CD117 [21], forward 5'-AACGACACGCTGGTCCGCTG-3' and reverse 5'-GTACACAGAACTAGACACATC-3' giving a PCR product of 341 bp; CD24 [22], forward 5'-TGCTCCTACCCACGCAGATT-3' and reverse 5'-GGCCAACCCAGAGTTGGAA-3' giving a PCR product of 88 bp; CD29 [23], forward 5'-GTTACACGGCTGCTGGTCTT-3' and reverse 5'-CTACTGCTGACTTAGGGATC-3' giving a PCR product of 264 bp; cyclophilin [24], forward 5'-CCACCGTGTTCTTCGACATC-3' and reverse 5'-GGTCCAGCATTTGCCATGG-3' giving a PCR product of 302 bp. PCR was performed using 0.4 µM of each primer, 200 µM dNTP, 1.5-3.0 mM $MgCl_2$, 5 U AmpliTaq DNA polymerase (Applied Biosystems, Foster City, CA, USA) and approximately 2 µg template cDNA in a 1X PCR buffer (50 mM KCl, 10 mM Tris-HCl pH 8.3, 0.001%

gelatin). Amplifications were performed in a thermocycler using 41 cycles of 95°C for 30 seconds, 55-60°C for 30 seconds, 72°C for 1 minute, followed by a final extension at 72°C for 10 minutes. Amplification products were visualized under ultraviolet light after gel electrophoresis on a 2% agarose gel and staining with GelGreen Nucleic Acid Gel Stain (Botium, Hayward, CA, USA) or ethidium bromide.

Immunocytochemistry and confocal microscopy.

Umbilical cord-derived MNC were spread and dried onto glass slides. The cells were fixed in 4% paraformaldehyde for 10 minutes at room temperature. After washing in 1X phosphate buffer saline (PBS), the cells were blocked with 10% donkey serum (Millipore, Billerica, MA, USA) in PBS for 30 minutes at room temperature, followed by incubation in the following primary antibodies in blocking solution overnight at 4°C: mouse anti-EpCAM (Clone E144, 1:100 dilution; Abcam, Cambridge, MA, USA) and rat anti-CD49f (Clone NKI-GoH3, 1:200 dilution; Millipore); rat anti-CD49f and mouse anti-CD24 (Clone SN3, 1:200 dilution; Millipore); and rabbit anti-CD29 (Clone EP1041Y, 1:200 dilution; Millipore) and mouse anti-CD24. The cells were then washed and incubated with the appropriate secondary antibodies for 1 hour at room temperature in the dark: Alexa Fluor 488 donkey anti-mouse, Alexa Fluor 488 donkey anti-rabbit, Alexa Fluor 594 donkey anti-rat, and/or Alexa Fluor 568 donkey anti-rabbit (all 1:200 dilution; Invitrogen/Molecular Probes, Eugene, OR, USA). After washing, the cells were stained in DRAQ5 (Cell Signaling Technology, Danvers, MA, USA) and cover-slipped in Prolong Anti-fade reagent (Invitrogen/Molecular Probes). Primary antibodies were omitted for negative controls. Cellular co-localizations were examined with a True Confocal

Scanning Spectrophotometer microscope (Leica Microsystems, Deerfield, IL, USA) using excitation wavelengths 488 nm for Alexa Fluor 488; 568 nm for Alexa Fluor 568 and Alexa Fluor 594; and 633 nm for DRAQ5. Optical scanning was performed every 0.5 µm of cell thickness by a sequential scanning method.

Flow cytometric analyses.

Flow cytometric analyses were performed as described previously [5]. Briefly, 1 X 10^6 umbilical cord blood-derived MNC were incubated for 30 minutes on ice in the dark with the following fluorochrome-conjugated antibodies: anti-CD34-fluorescein isothiocyanate (FITC; Clone 581; BD BioSciences Pharmingen, San Diego, CA, USA), anti-CD38-phycoerythrin (PE; Clone HIT2; BD Biosciences Pharmingen), anti-EpCAM-FITC (Clone VU-1D9; StemCell Technologies, Vancouver, Canada), anti-CD49f-PE (Clone GoH3; BD BioSciences Pharmingen), anti-CD117-allophycocyanin (APC; Clone YB5.B8; BD BioSciences Pharmingen), anti-CD24-FITC (Clone SN3; Antibodies-online, Aachen, Germany), anti-CD29-APC (Clone MAR4; BD BioSciences Pharmingen), or the combination of anti-CD34-FITC and anti-CD38-PE, or the combination of anti-EpCAM-FITC, anti-CD49f-PE, and anti-CD117-APC, or the combination of anti-CD24-FITC, anti-CD49f-PE and anti-CD29-APC. Samples treated with no antibody served as negative controls. Cells were washed, fixed with 4% paraformaldehyde and analyzed using a FACSCalibur flow cytometer (BD Biosciences Immunocytometry Systems, San Jose, CA, USA). Hematopoietic ($CD34^+$, $CD34^+CD38^-$) and putative breast ($EpCAM^+$, $EpCAM^+CD49f^+$, $EpCAM^+CD49f^+CD117^+$, $CD49f^+CD24^+$, $CD24^+CD29^+$, $CD49f^+CD24^+CD29^+$) stem/progenitor cell subpopulations were quantified from the

gated MNC population (lymphocytes and monocytes based on forward versus side light scatter) using the FlowJo software program (Tree Star, Ashland, OR, USA). The number of cells in these populations was normalized to 10^3 MNC.

Statistical analysis.

Descriptive statistics on the characteristics of study population and laboratory data were summarized. Spearman rank correlation coefficients were estimated for bivariate analyses between levels of different stem cell subpopulation. Statistical significance was set at 0.05 (two-sided). STATA version 11 (StataCorp LP, College Station, TX, USA) was used to conduct statistical analyses.

RESULTS

To determine whether putative breast stem/progenitor cell phenotypes were present in umbilical cord blood, we analyzed genes reported for breast stem cell markers [19] in the MNC fraction of umbilical cord blood by RT-PCR. Because breast stem cells are considered epithelial in nature [3], we first detected the gene for epithelial adhesion molecule (EpCAM), or epithelial-specific antigen (ESA), as a marker for epithelial cells [25]. Additionally, genes of putative markers for breast stem/progenitor cells, i.e., CD49f (α6-integrin), CD117 (c-kit receptor), CD24, and CD29 (β1-integrin) were detected in umbilical cord blood-derived MNC (Figure 1A).

Second, we determined protein expressions by immunocytochemistry and observed co-localized staining of $EpCAM^+CD49f^+$, $CD49f^+CD24^+$ and $CD24^+CD29^+$ surface markers in umbilical cord blood-derived MNC by confocal microscopic analyses

(Figure 1B). We further quantified the percentages of umbilical cord blood-derived MNC with putative markers of the different breast stem cell subpopulations ($EpCAM^+CD49f^+$, $EpCAM^+CD49f^+CD117^+$, $CD49f^+CD24^+$, $CD24^+CD29^+$, and $CD24^+CD29^+CD49f^+$) in addition to EpCAM by flow cytometry (Figure 1C). Data analyses using the FlowJo software program showed that the $EpCAM^+$ subpopulation ranged from 0.19 to 19.8 cells/1,000 MNC with a mean of 3.4 ± 4.0 cells; the $EpCAM^+CD49f^+$ subpopulation ranged from 0.049 to 9.7 cells/1,000 MNC with a mean of 1.7 ± 1.8 cells; the $EpCAM^+CD49f^+CD117^+$ subpopulation ranged from 0.02 to 2.4 cells/1,000 MNC with a mean of 0.48 ± 0.56 cells; the $CD49f^+CD24^+$ subpopulation ranged from 0 to 48.1 cells/1,000 MNC with a mean of 14.7 ± 12.9 cells; the $CD24^+CD29^+$ subpopulation ranged from 0.11 to 46.2 cells/1,000 MNC with a mean of 10.3 ± 9.8 cells; and the $CD24^+CD29^+CD49f^+$ subpopulation ranged from 0 to 44.4 cells/1,000 MNC with a mean of 8.3 ± 8.8 cells (Table 1).

We also quantified the percentages of umbilical cord blood-derived MNC with hematopoietic stem cell markers, i.e., $CD34^+$ and $CD34^+CD38^-$, and performed a rank correlation analysis between concentrations of hematopoietic and breast stem/progenitor cell subpopulations. Levels of the $EpCAM^+$ subpopulation were positively correlated with concentrations of $CD34^+$ and $CD34^+CD38^-$ hematopoietic stem cells (both $P = 0.006$; Table 2). Except for the $CD24^+CD29^+$ cells, all putative breast stem/progenitor cell subpopulations were positively associated with the hematopoietic stem cell subpopulations. Notably, the $EpCAM^+CD49f^+$ and $CD49f^+CD24^+$ subpopulations were positively and significantly correlated to the $CD34^+$ cells ($P = 0.03$

and 0.008, respectively). These associations were clearer among female than among male newborns, in particular the $EpCAM^+$ subpopulation (Table 2).

DISCUSSION

To our knowledge, this is the first report of measurable breast stem cells in human umbilical cord blood. Levels of these breast stem cells correlated significantly and positively with that of hematopoietic stem cells; concentrations of hematopoietic stem cells (as surrogate measurements of overall stem cell levels in the intrauterine environment) are correlated with umbilical cord blood plasma levels of IGF-1 and with birth weight [4-6].

Because the procurement of human cord blood samples is unpredictable, we chose to assay for putative breast stem/progenitor cells by analyzing published surface markers instead of using live cell-based methods employing dyes such as Hoechst 33342 or ALDEFLUOR [26, 27]. Since all cells within the mammary epithelium - except for the myoepithelial cells - express EpCAM, or ESA [28, 29] and breast stem cells are considered to be epithelial in nature [3, 30], we initially examined and found the gene for EpCAM - a marker of many epithelial cells [25] - in the MNC of umbilical cord blood. In fresh, uncultured umbilical cord blood-derived MNC, we also detected other putative genes related to breast stem/progenitor cells; these include CD117 (c-kit), CD24, and CD29 (β1-integrin), the latter two markers originally reported in mouse tissues [19, 31-33].

By immuno-staining umbilical cord blood-derived MNC, we demonstrated the co-expression of EpCAM and CD49f (α6-integrin), a major cellular phenotype of human

breast stem/progenitor cells [28, 34, 35]. The co-expressions of the proteins for CD24 and CD49f, and CD24 and CD29, indicate that mouse and human do share common markers of putative breast stem/progenitor cells.

The detection of putative breast stem/progenitor cell phenotypes suggests that there is a “mammary” compartment within the umbilical cord blood. These putative breast stem/progenitor cell phenotypes in the umbilical cord blood niche could be rare, but detectable by flow cytometry (Figure 1C, Table 1). Our results indicate that, at the time of birth, intrauterine conditions may sustain certain subpopulations of breast stem/progenitor cell phenotypes. It is not clear, however, how these breast phenotypes come about. Possibly, umbilical cord blood cells with breast phenotypes were derived from embryonic-like stem cells [14, 36-40] due to exposures to specific hormones or growth factors *in utero*, or dedifferentiated from hematopoietic stem cells as a result of epigenetic reprogramming, as proposed for the neural phenotypes found in umbilical cord blood [41]. These mammary phenotypes might also be differentiated from multipotent mesenchymal stem cells present in umbilical cord blood [18].

The concentration of umbilical cord blood-derived MNC carrying the epithelial EpCAM antigen was positively associated with the concentrations of hematopoietic stem cell populations identified by the CD34 surface marker [42, 43] and the subpopulation that was positive for CD34 but negative for CD38 [44]. Notably, the $EpCAM^{+}CD49f^{+}$ and $CD49f^{+}CD24^{+}$ breast stem/progenitor cell phenotypes showed a positive correlation with the $CD34^{+}$ hematopoietic stem cells. This finding supports the “stem cell burden” hypothesis [3] in predicting breast cancer risk. Non-significant associations might be due to the limited sample sizes for markers reported more

recently and thus incorporated in later stages of our study. The inconsistent associations observed for the strictly "mouse" $CD24^+CD29^+$ subpopulation suggest that such cells might not be functional mammary phenotypes in humans, although the $CD49f^+CD24^+$ subpopulation was originally reported also in mice [32]. These correlations should be re-examined in larger sample sizes, as should some of the significant associations observed in the analyses stratified by gender.

In summary, early life exposures have been linked to risk of breast cancer in the offspring through a pathway hypothesized to involve the mammary stem cell pool. The "breast stem cell burden and susceptibility" hypothesis is based on the assumption that breast cancer originates from mammary stem cells [3, 9, 10, 45]. The greater the number of breast stem cells, the higher the likelihood that one of these cells will undergo malignant transformation. Because mammary stem cells arise primarily during the fetal/perinatal period, the *in utero* environment becomes a major determinant of their number. Hence, a breast "stem cell potential" – a term proposed for measurable variables that reflect the effects of intrauterine and perinatal influences on stem cell burden and susceptibility - in umbilical cord blood might predict subsequent risk of breast cancer in the adult life. Future research will determine whether these putative populations with breast stem/progenitor phenotypes in umbilical cord blood are indeed functional mammary cells. If so, we will have a model system to understand whether genetic and/or epigenetic alterations in breast stem/progenitor cells explains fetal programming of breast cancer risk in the adult life.

Funding

This work was supported by an Innovator Award from the US Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Program [W81XWH-05-1-0314 to D.T.] and a grant from the US National Institute of Health [CA090902 to C.-C.H.].

Acknowledgements

We gratefully acknowledge the contribution of the late Dr. Todd Savarese during the early stages of the study. We thank Chandni Amaresan and Kimberly Cheng for their assistance during their internships at the laboratory.

Disclosure

The authors have declared no conflicts of interest.

REFERENCES

1. Trichopoulos D. Hypothesis: does breast cancer originate in utero? Lancet 1990; 335: 939-940.
2. Baik I, Becker PS, DeVito WJ et al. Stem cells and prenatal origin of breast cancer. Cancer Causes Cont 2004; 15: 517-530.
3. Savarese TM, Low HP, Strohsnitter WC, Hsieh CC. Normal breast stem cells, malignant breast stem cells and the perinatal origin of breast cancer. Stem Cell Rev 2006; 2: 103-110.
4. Baik I, DeVito WJ, Ballen K et al. Association of fetal hormone levels with stem cell potential: evidence for early life roots of human cancer. Cancer Res 2005; 65; 358-363.
5. Savarese TM, Strohsnitter WC, Low HP et al. Correlation of umbilical cord blood hormones and growth factors with stem cell potential: implications for the prenatal origin of breast cancer hypothesis. Breast Cancer Res 2007; 9: R29.
6. Strohsnitter WC, Savarese TM, Low HP et al. Correlation of umbilical cord blood hematopoietic stem and progenitor cell levels with birth weight: implications for prenatal influence on cancer risk. Br J Cancer 2008; 98: 660-663.
7. Michels KB, Xue F. Role of birthweight in the etiology of breast cancer. Int J Cancer 2006; 119: 1007-1025.
8. Park SK, Kang D, McGlynn KA et al. Intrauterine environments and breast cancer risk: meta-analysis and systematic review. Breast Cancer Res 2008; 10: R8.

9. Trichopoulos D, Lagiou P, Adami HO. Towards an integrated model for breast cancer etiology: the crucial role of the number of mammary tissue-specific stem cells. Breast Cancer Res. 2005; 7: 13-17.

10. Trichopoulos D, Adami HO, Ekbom A et al. Early life events and conditions and breast cancer risk: From epidemiology to etiology. Int J Cancer. 2008; 122: 481-485.

11. Mayani H, Lansdorp PM. Biology of human umbilical cord blood-derived hematopoietic stem/progenitor cells. Stem Cells 1998; 16: 153-165.

12. Hill JM, Zalos G, Halcox JPJ et al. Circulating endothelial progenitor cells, vascular function, and cardiovascular risk. N Engl J Med 2003; 348: 593-600.

13. Ingram DA, Mead LE, Tanaka H et al. Identification of a novel hierarchy of endothelial progenitor cells using human peripheral and umbilical cord blood. Blood 2004; 104: 2752-2760.

14. Zuba-Surma EK, Klich I, Greco N et al. Optimization of isolation and further characterization of umbilical cord blood-derived very small embryonic/epiblast-like stem cells (VSELs). Euro J Haematol 2009; 84: 34-46.

15. Ha Y, Lee JE, Kim KN et al. Intermediate filament nestin expressions in human cord blood monocytes (HCMNCs). Acta Neurochir 2003; 145: 483-487.

16. Zangiacomi V, Balon N, Maddens S et al. Human cord blood-derived hematopoietic and neural-like stem/progenitor cells are attracted by the neurotransmitter GABA. Stem Cells Dev 2009; 18: 1369-1377.

17. Kögler G, Sensken S, Airey JA et al. A new human somatic stem cell from placental cord blood with intrinsic pluripotent differentiation potential. J Exp Med 2004; 200: 123-135.

18. Wagner W, Wein F, Seckinger A et al. Comparative characteristics of mesenchymal stem cells from human bone marrow, adipose tissue, and umbilical cord blood. Exp Hematol 2005; 33: 1402-1416.

19. Visvader JE. Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev 2009; 23: 2563-2577.

20. Gaudreault M, Vigneault F, Gingras ME et al. Transcriptional regulation of the human α6 integrin gene by the transcription factor NFI during corneal wound healing. Invest Ophthalmol Vis Sci 2008; 49: 3758-3767.

21. Rassidakis GZ, Georgakis GV, Oyarzo M et al. Lack of c-kit (CD117) expression in CD30+ lymphomas and lymphomatoid papulosis. Modern Pathol 2004; 17: 946-953.

22. Kaipparettu BA, Malik S, Konduri SD et al. Estrogen-mediated downregulation of CD24 in breast cancer cells. Int J Cancer 2008; 123: 66-72.

23. Bouchet BY, Colón M, Polotsky A et al. Beta-1 integrin expression by human nasal chondrocytes in microcarrier spinner culture. J Biomed Mater Res 2000; 52: 716-724.

24. Mayer-Proschel M, Kalyani AJ, Mujtaba T, Rao MS. Isolation of lineage-restricted neuronal precursors from multipotent neuroepithelial stem cells. Neuron 1997; 19: 773-785.

25. Litvinov SV, Velders MP, Bakker HAM et al. Ep-CAM: a human epithelial antigen is a homophilic cell-cell adhesion molecule. J Cell Biol 1994; 125: 437-446.

26. Alvi AJ, Clayton H, Joshi C et al. Functional and molecular characterization of mammary side population cells. Breast Cancer Res 2003; 5: R1-R8.

27. Ginestier C, Hur MH, Charafe-Jauffret et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell 2007; 1: 555-567.

28. Stingl J, Eaves CJ, Zandieh I, Emerman JT. Characterizaton of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. Breast Cancer Res Treat 2001; 67: 93-109.

29. Gudjonsson T, Villadsen R, Nielsen HL et al. Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. Genes Dev 2002; 16: 693-706.

30. Smith GH, Chepko G. Mammary epithelial stem cells. Microsc Res Tech 2001; 52: 190-203.

31. Shackleton M, Vaillant F, Simpson KJ et al. Generation of a functional mammary gland from a single stem cell. Nature 2006; 439: 84-88.

32. Stingl J, Eirew P, Ricketson I et al. Purification and unique properties of mammary epithelial stem cells. Nature 2006; 439: 993-997.

33. Visvader JE, Lindeman GJ. Mammary stem cells and mammopoiesis. Cancer Res 2006; 66: 9798-9801.

34. Stingl J, Raouf A, Eirew P, Eaves CJ. Deciphering the mammary epithelial cell hierarchy. Cell Cycle 2006; 5: 1519-1522.

35. Villadsen R, Fridriksdottir AJ, Rønnov-Jessen L et al. Evidence for a stem cell hierarchy in the adult human breast. J Cell Biol 2007; 177: 87-101.

36. Baal N, Reisinger K, Jahr H et al. Expression of transcription factor Oct-4 and other embryonic genes in CD133 positive cells from human umbilical cord blood. Thromb Haemost 2004; 92: 767-775.

37. McGuckin CP, Forraz N, Baradez MO et al. Production of stem cells with embryonic characteristics from human umbilical cord blood. Cell Proliferation 2005; 38: 245-255.

38. Habich A, Jurga M, Markiewicz I et al. Early appearance of stem/progenitor cells with neural-like characteristics in human cord blood mononuclear fraction cultured in vitro. Exp Hematol 2006; 34: 914-925.

39. Zhao Y, Wang H, Mazzone T. Identification of stem cells from human umbilical cord blood wuth embryonic and hematopoietic characteristics. Exp Cell Res 2006; 312: 2454-2464.

40. Halasa M, Baskiewicz-Masiuk M, Dabkowska E, Machalinski B. An efficient two-step method to purify very small embryonic-like (VSEL) stem cells from umbilical cord blood (UCB). Folia Histochem Cytobiol 2008; 46: 239-243.

41. Domanska-Janik K, Buzanska L, Lukomsja B. A novel, neural potential of non-hematopoietic human umbilical cord blood stem cells. Int J Dev Biol 2008; 52: 237-248.

42. Civin CI, Strauss LC, Brovall C et al. Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells. J Immunol 1984; 133: 157-165.

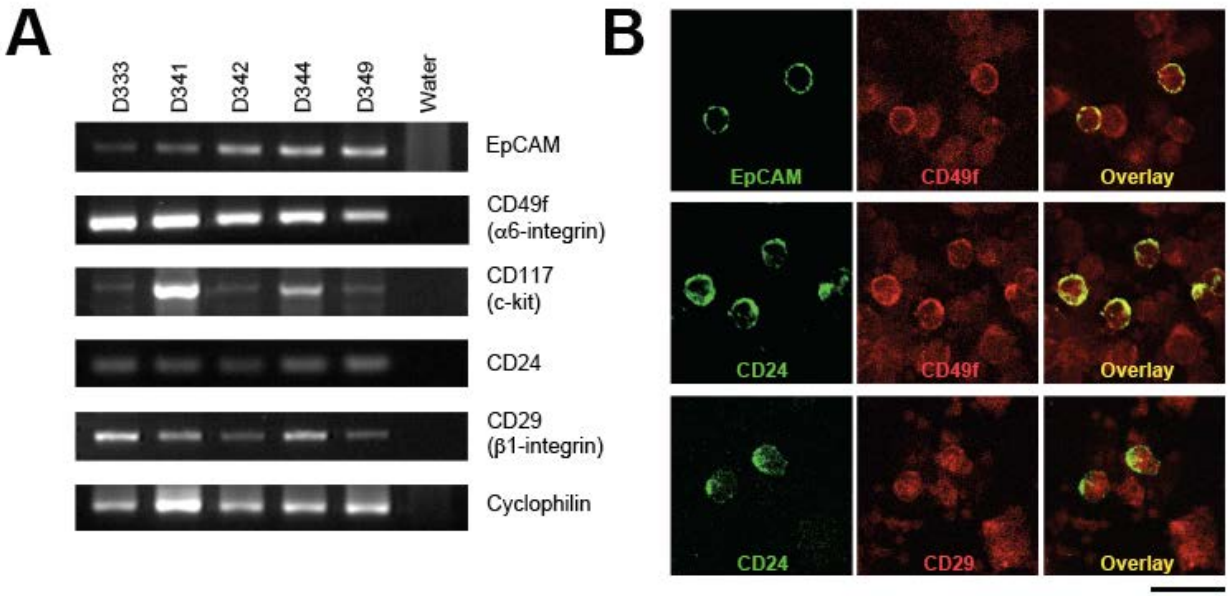
43. Sutherland DR, Stewart AK, Keating A. CD34 antigen: molecular features and potential clinical applications. Stem Cells 1993; 11: 50-57.

44. Xiao M, Dooley DC. Cellular and molecular aspects of human $CD34^{+}CD38^{-}$ precursors: analysis of a primitive hematopoietic population. Leuk Lymphoma 2000; 38: 489-497.

45. Ginestier C, Wicha MS. Mammary stem cell number as a determinate of breast cancer risk. Breast Cancer Res 2007; 9: 109-110.

FIGURE LEGENDS

Figure 1. (A) Gel electrophoresis showing the detection of PCR products for EpCAM, CD49f (α6-integrin), CD117 (c-kit), CD24, CD29 (β1-integrin), and the housekeeping gene cyclophilin from umbilical cord blood-derived mononuclear cells from 5 umbilical cord blood samples (D333, D341, D342, D344, and D349). Water (last lane) was used as negative controls. (B) Double-labeled immunofluorescent confocal microscopy of umbilical cord blood-derived mononuclear cells from sample D321 showing colocalization in the overlay image of EpCAM (green) and CD49f (red) (top panel); CD24 (green) and CD49f (red) (middle panel); and CD24 (green) and CD29 (red) (bottom panel). Scale bar represents 20 µm. (C) Flow cytometric pseudocolor plots showing the detection of the $EpCAM^+$, $EpCAM^+CD49f^+$, $EpCAM^+CD49f^+CD117^+$, $CD49f^+CD24^+$, $CD24^+CD29^+$, and $CD49f^+CD24^+CD29^+$ subpopulations (boxed, with percentage of cells indicated) from umbilical cord blood-derived mononuclear cells of sample N41. The arrows indicate that the triple positive population was derived from the double positive population as shown. The markers shown in the top panel have been reported in humans while the markers shown in the bottom panel have been reported in mice.


A
D333
D341
D342
D344
D349
Water
EpCAM
CD49f
(α6-integrin)
CD117
(c-kit)
CD24
CD29
(β1-integrin)
Cyclophilin
B
EpCAM
CD49f
Overlay
CD24
CD49f
Overlay
CD24
CD29
Overlay

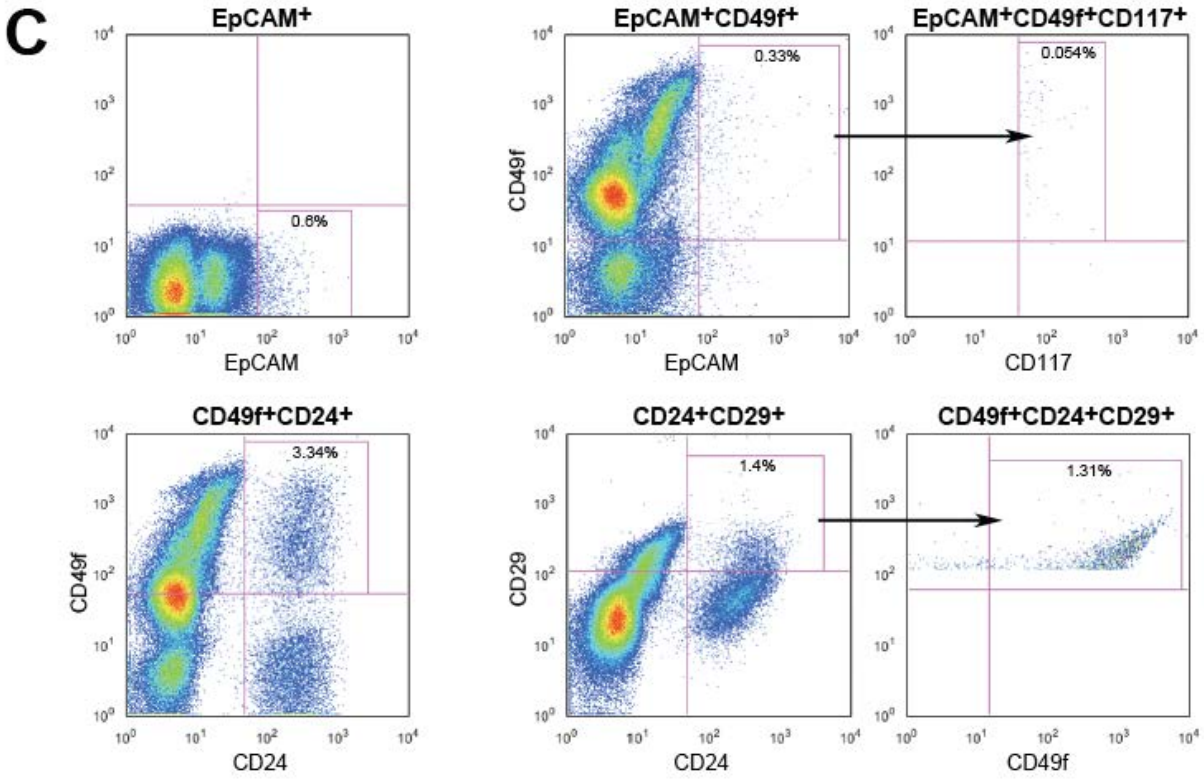

C
EpCAM+
0.6%
EpCAM
EpCAM+CD49f+
0.33%
CD49f
EpCAM
EpCAM+CD49f+CD117+
0.054%
CD117
CD49f+CD24+
3.34%
CD49f
CD24
CD24+CD29+
1.4%
CD29
CD24
CD49f+CD24+CD29+
1.31%
CD49f

Table 1. Maternal and newborn characteristics, and umbilical cord blood stem/progenitor cell counts.

Subject characteristics	*N*	Mean ± SD or %	Range
Mother's age (yr)	169	30.1 ± 6.2	19 - 44
Parity			
First	52	30.8	
Second	49	29.0	
Third	24	14.2	
Fourth and above	44	26.0	
Gestation Duration (wks)	170	39.2 ± 1.2	37 - 41
Newborn gender			
Male	84	49.4	
Female	86	50.6	
Newborn birth weight (g)	170	3,363.7 ± 510.5	1,973 - 4,917
Umbilical cord blood volume (ml)[a]	169	98.4 ± 29.4	45 - 216
Umbilical cord blood stem/progenitor cell populations[b]			
$CD34^+$	169	8.5 ± 7.4	0.0 - 56.7
$CD34^+CD38^-$	167	2.4 ± 2.5	0.0 - 23.8
$EpCAM^+$	112	3.4 ± 4.0	0.19 - 19.8
$EpCAM^+CD49f^+$	109	1.7 ± 1.8	0.049 - 9.7
$EpCAM^+CD49f^+CD117^+$	32	0.48 ± 0.56	0.02 - 2.4
$CD49f^+CD24^+$	56	14.7 ± 12.9	0.0 - 48.1
$CD24^+CD29^+$	55	10.3 ± 9.8	0.11 - 46.2
$CD49f^+CD24^+CD29^+$	55	8.3 ± 8.8	0.0 - 44.4

[a]Includes 35 ml of citrate-phosphate-dextrose anticoagulant.
[b]Cell counts per 10^3 MNC.

Table 2. Spearman correlation coefficients (*P* values in parentheses) between umbilical cord blood hematopoietic and breast stem/progenitor cell populations.

Hematopoietic stem cell subpopulations	Breast stem/progenitor cell subpopulations					
	$EpCAM^+$	$EpCAM^+CD49f^+$	$EpCAM^+CD49f^+CD117^+$	$CD49f^+CD24^+$	$CD24^+CD29^+$	$CD49f^+CD24^+CD29^+$
All subjects (*N*)	112	109	32	56	55	55
$CD34^+$	0.26 (**0.006**)	0.21 (**0.03**)	0.24 (0.20)	0.35 (**0.008**)	0.14 (0.32)	0.17 (0.21)
$CD34^+CD38^-$	0.26 (**0.006**)	0.15 (0.12)	0.30 (0.09)	0.24 (0.07)	0.01 (0.97)	0.18 (0.19)
Newborn gender						
Males (*N*)	53	53	14	25	25	25
$CD34^+$	0.23 (0.37)	0.22 (0.12)	0.41 (0.15)	0.33 (0.11)	0.07(0.76)	0.28 (0.18)
$CD34^+CD38^-$	0.02 (0.90)	0.04 (0.77)	0.27 (0.36)	0.38 (0.06)	0.09 (0.67)	0.36 (0.08)
Females (*N*)	59	56	18	31	30	30
$CD34^+$	0.36 (**0.005**)	0.18 (0.19)	0.03 (0.90)	0.35 (**0.05**)	0.18 (0.34)	0.05 (0.78)
$CD34^+CD38^-$	0.37 (**0.004**)	0.19 (0.17)	0.18 (0.48)	0.19 (0.29)	-0.03 (0.88)	0.08 (0.68)

APPENDIX 3

PDF copies of perspectives / commentaries on the early life origins of breast cancer supported by the W81XWH-05-1-0314 Innovator Award (activities in the context of CP6)

Int. J. Cancer: **122,** 481 485 (2008)

MINI REVIEW

Early life events and conditions and breast cancer risk: From epidemiology to etiology

Dimitrios Trichopoulos[1]*, Hans-Olov Adami[1,2], Anders Ekbom[3], Chung-Cheng Hsieh[1,4] and Pagona Lagiou[1,5]

[1]*Department of Epidemiology, Harvard School of Public Health, Boston, MA*
[2]*Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden*
[3]*Department of Medicine, Unit of Clinical Epidemiology, Karolinska Institutet/Karolinska University Hospital, Stockholm, Sweden*
[4]*Department of Cancer Biology and Cancer Center, University of Massachusetts Medical School, Worcester, MA*
[5]*Department of Hygiene and Epidemiology, School of Medicine, University of Athens, Greece*

Risk factors for breast cancer documented by intensive epidemiological investigations and viewed in the context of general principles of carcinogenesis can be integrated to an etiologic model comprising 3 principal components: the likelihood of breast cancer occurrence depends on the number of mammary tissue-specific stem cells, which is determined in early life; all growth-enhancing mammotropic hormones affect the rate of expansion of initiated clones; and while a pregnancy stimulates the replication of already initiated cells, it conveys long-term protection through differentiation of mammary tissue-specific stem cells. This perspective accommodates much of what is known about the epidemiology and natural history of breast cancer and highlights the role of early life in the origin of this cancer.

Key words: breast cancer; perinatal; birth weight; stem cells; hormones; pregnancy

Breast cancer epidemiology

The incidence of breast cancer has apparently increased throughout the world during the last century, even before the widespread application of mammographic screening programs and mortality from the disease in developed countries generally exceeds that from other cancer sites.[1] Breast cancer epidemiology has been intensively studied, perhaps more than that of any other cancer.[2 6] Table I summarizes what are generally considered as established epidemiological characteristics of breast cancer and provides an indication of the strength of the respective associations in terms of the relative risk per natural contrasts or usual increments.

Breast cancer is mostly, though not exclusively, a disease of women. The incidence of the disease increases with age, with an inflection around menopause, which is not evident for other forms of cancer. It is generally more common among urban rather than rural residents as well as among women of higher socioeconomic status. In comparison to Asian women in China or Japan, Caucasian women in the western world have a considerably higher breast cancer risk.[1]

With respect to reproductive history, an earlier age at menarche and a later age at menopause are associated with increased risk whereas, for a given age at menopause, induced menopause conveys more protection than the naturally occurring one.[6 8] The role of pregnancies is complex. Irrespective of the woman's age, a pregnancy imparts a short term increase of breast cancer risk[9] followed by a substantial long term reduction of this risk, as was first documented with respect to the first pregnancy some 40 years ago in a classical international epidemiological study.[10] Hence, the earlier the age at first full term pregnancy, the more prolonged is the subsequent long term protection. After the age of about 35 years, a first pregnancy actually increases breast cancer risk, apparently because the short term risk increase exceeds the subsequent risk reduction. Additional full term pregnancies have similar but much weaker effects,[11] while spontaneous or induced abortions do not affect breast cancer risk.[12] Prolonged lactation conveys at most modest protection, which appears to be restricted to premenopausal women.[13,14] Current or recent use of oral contraceptives slightly increase the risk for breast cancer,[15] whereas long term use of replacement estrogens with progestins may substantially increase breast cancer risk.[16 18]

High birth weight has been associated with increased breast cancer risk in the offspring.[19] Having been breastfed as an infant has been investigated for its role in breast cancer under the assumption that it could be responsible for the transmission of an infectious agent, but the results did not support an association.[20] Early life growth[21] and factors that may increase it[22] have also been positively associated with breast cancer risk, as has height[23,24] and post (but not pre) menopausal obesity[8,25 27] later in life.

A high density mammogram (75% or more of the total breast area with dense mammographic appearance) has been associated with a more than 4 fold risk in comparison to a low density mammogram (10% or less or total breast area with dense mammographic appearance).[28] Atypical hyperplasia of the mammary gland has been documented as an important breast cancer risk factor.[29,30]

Family history among first degree relatives is associated with increased breast cancer risk.[31] BRCA1 and BRCA2, as well as some highly penetrant mutations, explain a large part of familial breast cancers, but account for a small proportion of all breast cancers.[32] Many studies have examined low penetrance susceptibility polymorphisms in candidate genes, but the associations reported in some studies could not been replicated in subsequent investigations. This is an evolving field, in which large whole genome association investigations are providing new insight.[33] Breast cancer in the contralateral breast is an established risk factor for developing the disease in the other breast, but the underlying pathogenetic mechanisms are not clear.[34]

High levels of physical activity[35] and high intake of vegetables, perhaps fruits[36] and olive oil[37] have been reported to be associated with reduced breast cancer risk, possibly by reducing endogenous estrogen levels.[38,39] Nevertheless, the evidence is inconclusive and suggests, at most, weak effects. Recent evidence points to total and particularly saturated fat as been weakly, but significantly, positively associated with breast cancer risk.[40] Most studies indicate that consumption of alcoholic beverages may slightly

---

Based on the Third Richard Doll lecture given at the International Agency for Research on Cancer by DT on May 9, 2007 and relying on the collective work of all authors.

Grant sponsors: W81XWH 05 1 0314 Innovator Award, US. Department of Defence Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs.

***Correspondence to:** Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. Fax: +1 617 566 7805. E mail: dtrichop@hsph.harvard.edu

Received 24 October 2007; Accepted after revision 25 October 2007
DOI 10.1002/ijc.23303
Published online 16 November 2007 in Wiley InterScience (www.interscience.wiley.com).

TABLE I – FACTORS EVALUATED IN RELATION TO BREAST CANCER RISK

Risk factor	Category/change	Strength
Gender	Women vs. men	++++
Age	Increase	++++
Ethnic group	Caucasian vs. Asian	+++
Family history	Yes vs. no	+++
Specific genes	Yes vs. no	++++
Cancer in other breast	Yes vs. no	+++
Height	Increase	++
Postmenopausal obesity	Increase	++
Birth weight	Increase	+
Having been breastfed	No vs. yes	0
Growth in early life	Increase	+
Atypical hyperplasia	Present vs. absent	+++
Mammographic density (mammary gland mass)	High vs. low density (increasing mass)	+++
Age at menarche	Earlier	++
Age at menopause	Later	++
Type of menopause	Natural vs. artificial	++
Age at 1st full term pregnancy	Later	+++
Age at other pregnancies	Later	+
Parity overall	Lower	++
Pregnancy timing	Proximal vs. distant	+
Lactation	No vs. yes	+
Abortion	No vs. yes	0
Oral contraceptive use (recent)	Increase	+
Hormone replacement	Increase	++
Plant foods and olive oil	Reduced intake	+
Saturated fat	Increased intake	+
Physical activity	Reduced	+
Ethanol intake	Increase	+
Ionizing radiation	Increased	+
Magnetic fields	Increased	0
Organochlorines	Increased	0

Association: ++++ very strong, +++ strong, ++ modest, + weak, 0 null.

increase breast cancer risk, possibly by increasing estrogen levels.[41] There is no conclusive evidence for an association between tobacco smoking and the disease.[42,43]

Ionizing radiation is an established cause of cancer of the breast as well as of several other cancers, but it is of limited quantitative importance. Exposure to organochlorines[44] or electromagnetic fields[45] has not been shown to be related to breast cancer.

It is generally believed that the association between endogenous hormones and breast cancer risk should be studied in prospective, rather than retrospective, investigations, under the undocumented assumption that disease status, even prior to treatment, may affect hormone levels. Among postmenopausal women, most hormones examined with the notable exception of adiponectin that has been mostly evaluated through case control designs[46,47] have been positively associated with breast cancer risk.[6,48–50] The list includes total and free estradiol, estrone and estrone sulphate, androstenedione, dehydroepiandrosterone and dehydroepiandrosterone sulphate, testosterone and prolactin. Among premenopausal women, case control studies and a few cohort investigations provide some support for a positive association between estrogens and breast cancer risk, but they also indicate that high levels of androgens could increase this risk.[6,51] In both prospective and retrospective studies among premenopausal women, significant positive associations have been reported between blood insulin like growth factor 1 (IGF 1) and breast cancer risk.[52]

The early life etiological model

The etiological model we have proposed for breast cancer accommodates most of what we know about the epidemiology of the disease. The model emphasizes early life events and conditions as determinants of breast cancer risk and summarizes the distinct epidemiological characteristics of the disease on the basis of 3 major components[5,22,53–58]:

- The likelihood of breast cancer occurrence depends on the number of mammary tissue specific stem cells, which is determined early in life, including the intrauterine life,
- in early and later life, growth enhancing mammotropic hormones affect the replication rate of mammary tissue specific stem cells, the likelihood of retention of cells with spontaneous somatic mutations as well as the rate of expansion of initiated clones, and
- while a pregnancy stimulates the replication of already initiated cells, it conveys long term protection through differentiation of a large fraction of the mammary tissue specific stem cells.

It should be noted that several scientists[23,59,60] have postulated, explicitly or implicitly, that early life influences may play a role in breast cancer etiology and there have even been early studies exploring birth weight as a breast cancer risk factor.[61] Moreover, the issue of pregnancy induced terminal differentiation of mammary gland has been championed by Russo and Russo.[62]

Categorization of breast cancer risk factors according to the 3 components of the early life etiological model

An etiological model should accommodate the epidemiological profile of the disease it aims to explain. In this context, we have categorized the established breast cancer risk factors according to the 3 components of the early life etiological model, taking also into account that certain breast cancer epidemiologic characteristics reflecting general principles of carcinogenesis relevant to many cancer sites (Table II). The empirical evidence in support of the categorization has been presented in detail in earlier publications[2,5] and is summarized below.

First component

Mammary gland mass, *as distinct from breast size*, is usually assessed through mammographic density and is an important breast cancer risk factor.[28] Mammary gland mass, which is likely to reflect the pool of mammary cells and be correlated with the number of mammary stem cells,[63,64] can also accommodate several breast cancer risk factors, including the higher incidence of the disease among Caucasian compared to Asian women and women of higher rather than lower socioeconomic class as well as the preponderance of breast cancer in the slightly larger left, rather than right, breast.[65] The positive associations of breast cancer risk with birth weight, growth in early life and adult height could also be explained in terms of mammary gland mass. Finally, at the extreme, the strikingly higher breast cancer risk among women than among men, even in later life when estrogen production is not substantially different between the 2 genders, is best explained by the correspondingly higher mammary gland mass among women than among men.

Second component

Most investigators agree that oestrogens in general, or specific categories of oestrogens, or prolactin, or other hormones, including IGF, are important in the etiology of breast cancer. Our view is that all growth enhancing and mammotropic hormones are involved in one or more stages in the long process leading to clinical breast cancer. An important issue that has not been sufficiently explored in empirical research is the way these hormones interact in the causation of the disease. A small study presented evidence that mammotropic hormones may act as permissive factors for breast cancer occurrence and that values of any one of these above a certain level may suffice for sustaining growth of a developing tumor[66] the finding is intriguing but requires confirmation in larger datasets. The second component of the etiologic model accommodates our knowledge about the role of reproductive factors in the etiology of the disease

TABLE II – GROUPING OF BREAST CANCER RISK FACTORS ACCORDING TO THE GENERAL PRINCIPLES OF CARCINOGENESIS AND THE POSTULATED PATHOGENIC PROCESS

General principles of carcinogenesis	Number of mammary tissue specific stem cells	Growth enhancing mammotropic hormones	Terminal differentiation
Age	Mammographic density (gland mass)	Gender	Age at 1st full term pregnancy
Ionizing radiation	Atypical hyperplasia	Age incidence pattern	Age at other pregnancies
Family history	Gender	Age at menarche	Parity overall
Specific genes	Birth weight	Age at menopause	Lactation
	Growth in early life	Type of menopause	
	Height	Oral contraceptives	
	Ethnic group	Hormone replacement	
		Pregnancy timing	
		Postmenopausal obesity	
		Ethanol intake	
		Physical activity	
		Adult life diet	

as well as that of alcohol drinking (which tends to increase oestrogen levels), physical activity and adult life diet.[2,5]

Third component

Terminal differentiation of the mammary gland takes place mostly after the occurrence of the first full term pregnancy, and to a lesser extent, after the occurrence of subsequent pregnancies and lactation.[67] The later the age at first full term pregnancy, the higher the number of already initiated cells and the more limited the protection conveyed by pregnancy. Beyond the age of 35 or so, the transient increase of breast cancer risk that accompanies a pregnancy (due to the effect on already initiated clones of the many fold increases of mammotropic and growth enhancing hormones) overshadows the protection conveyed by the terminal differentiation of immature mammary cells. The 3rd component of the etiologic model also accommodates what was largely thought to be an enigma, namely why breast cancer risk is higher among parous than among nulliparous women of premenopausal age.

The ecological challenges

One of the most challenging characteristics in breast cancer epidemiology is the sharp ecological contrast in breast cancer incidence between women in western Europe and North America and women in China and Japan,[1] which fades in Asian women migrating to the west after 2 or more generations. Neither reproductive nor dietary factors in adult life can explain the 4 fold difference in incidence observed in these populations,[8,68] nor can they explain the subsequent incidence assimilation. On the contrary, diet in early life could provide an explanation for the ecological contrast in the context of the early life etiological model: reduced energy intake in early life is associated with smaller body size in adult life and smaller body size constrains birth weight and subsequent development of offspring. Increased energy intake, on the other hand, facilitates growth and removes constrains on birth weight and eventual body size. This cycle tends to repeat over consecutive generations of Asian migrants in western countries and is associated with a gradual increase in body size and breast cancer incidence among them.[22,57]

The early life etiological model is not refuted by the fact that populations at low risk for breast cancer have higher levels of most pregnancy or possibly adult life hormones.[69] It is plausible that in striking ecological contrasts (*e.g.* between native Chinese and Caucasian populations), pregnancy growth hormones tend to increase in order to compensate for physically constrained fetal growth[58,70] and the perinataly programmed higher levels of these hormones could track through adult life.

Avenues of future research

Future research assessing the early life aspects of the etiology of breast cancer could follow many directions and some of them are outlined here. Hsieh and coworkers[65,71,72] are evaluating how pregnancy mammotropic and growth hormones affect cord blood stem cell populations. In their recent work, they reported that cord blood plasma levels of IGF 1 were strongly correlated with all the hematopoietic stem and progenitor concentrations examined, whereas estriol and insulin like growth factor binding protein 3 levels were positively and significantly correlated with some of these cell populations. Hilakivi Clarke and her coworkers have used rodent models to explore the ways through which diet and otherwise induced epigenetic changes in target genes might lead to strategies to prevent breast cancer.[73,74] Critically important results may also emerge from a unique follow up of women born to mothers who have taken diethylstilbestrol (DES) during their pregnancies. Two recent publications indicated that in utero DES exposure may substantially increase breast cancer risk in the offspring.[75,76] Moreover, it would be important to firmly document what has already been reported in previous publications, that is, that perinatal characteristics predictive of high breast cancer risk in adult life are also predictive of high breast cancer risk mammographic patterns.[77,78]

Conclusion

The early life etiologic model we have outlined accommodates the existing epidemiological evidence. Its 3 components refer to stages of a single biological process that points to the number of mammary tissue specific stem cells as a core determinant of breast cancer risk. The first component focuses on the perinatal period, when stem cells are generated. The second component concentrates on preinitiation and postinitiation growth factors that modulate the number of mammary stem cells at risk and the growth of the initiated clones. The third postulate explains how cells at risk are removed through terminal differentiation.

References

1. Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2002 cancer incidence, mortality and prevalence worldwide. IARC CancerBase No. 5, version 2.0. Lyon: IARC, 2004.
2. Adami H O, Signorello LB, Trichopoulos D. Towards an understanding of breast cancer etiology. Seminars in Cancer Biology 1998:8: 255 262.
3. Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. Lancet Oncol 2001;2:133 40.
4. Richie RC, Swanson JO. Breast cancer: a review of the literature. J Insur Med 2003;35:85 101.
5. Trichopoulos D, Lagiou P, Adami HO. Towards an integrated model for breast cancer etiology: the crucial role of the number of mammary tissue specific stem cells. Breast Cancer Res 2005;7:13 17.
6. Hankinson S, Tamimi R, Hunter D. Breast cancer. In: Adami HO, Hunter D, Trichopoulos D, eds. Textbook of cancer epidemiology. New York: Oxford University Press, in press.

7. Trichopoulos D, MacMahon B, Cole P. Menopause and breast cancer risk. J Natl Cancer Inst 1972;48:605–13.
8. Hsieh C c, Trichopoulos D, Katsouyanni K, Yuasa S. Age at menarche, age at menopause, height and obesity as risk factors for breast cancer: associations and interactions in an international case control study. Int J Cancer 1990;46:796–800.
9. Lambe M, Hsieh C c, Trichopoulos D, Ekbom A, Pavia M, Adami H O. Transient increase in the risk of breast cancer after giving birth. N Engl J Med 1994;331:5–9.
10. MacMahon B, Cole P, Lin TM, Lowe CR, Mirra AP, Ravnihar B, Salber EJ, Valaoras VG, Yuasa S. Age at first birth and breast cancer risk. Bull World Health Organ 1970;43:209–21.
11. Trichopoulos D, Hsieh C, MacMahon B, Lin T, Lowe C, Mirra A, Ravnihar B, Salber E, Valaoras V, Yuasa S. Age at any birth and breast cancer risk. Int J Cancer 1983;31:701–4.
12. Melbye M, Wohlfahrt J, Olsen JH, Frisch M, Westergaard T, Helweg Larsen K, Andersen PK. Induced abortion and the risk of breast cancer. N Engl J Med 1997;336:81–5.
13. MacMahon B, Lin TM, Lowe CR, Mirra AP, Ravnihar B, Salber EJ, Trichopoulos D, Valaoras VG, Yuasa S. Lactation and cancer of the breast. A summary of an international study. Bull World Health Organ 1970;42:185–94.
14. Newcomb PA, Storer BE, Longnecker MP, Mittendorf R, Greenberg ER, Clapp RW, Burke KP, Willett WC, MacMahon B. Lactation and a reduced risk of premenopausal breast cancer. N Engl J Med 1994;330:81–7.
15. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53297 women with breast cancer and 100239 women without breast cancer from 54 epidemiological studies. Lancet 1996;347:1713–27.
16. Bergkvist L, Adami HO, Persson I, Hoover R, Schairer C. The risk of breast cancer after estrogen and estrogen progestin replacement. N Engl J Med 1989;321:293–7.
17. Beral V and Million Women Study Collaborators. Breast cancer and hormone replacement therapy in the Million Women Study. Lancet 2003;362:419–27.
18. Chlebowski RT, Hendrix SL, Langer RD, Stefanick ML, Gass M, Lane D, Rodabough RJ, Gilligan MA, Cyr MG, Thomson CA, Khandekar J, Petrovitch H, McTiernan A, WHI Investigators. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women's Health Initiative Randomized Trial. JAMA 2003;289:3243–53.
19. Michels KB, Xue F. Role of birthweight in the etiology of breast cancer. Int J Cancer 2006;119:2007–25.
20. Titus Ernstoff L, Egan KM, Newcomb PA, Baron JA, Stampfer M, Greenberg ER, Cole BF, Ding J, Willett W, Trichopoulos D. Exposure to breast milk in infancy and adult breast cancer risk. J Natl Cancer Inst 1998;90:921–4.
21. Ahlgren M, Melbye M, Wohlfahrt J, Sørensen TI. Growth patterns and the risk of breast cancer in women. N Engl J Med 2004;351:1619–26.
22. Lagiou P, Adami HO, Trichopoulos D. Early life diet and the risk for adult breast cancer. Nutr Cancer 2007;56:158–61.
23. Valaoras V, MacMahon B, Trichopoulos D, Polychronopoulou A. Lactation and reproductive histories of breast cancer patients in Greater Athens, 1965–1967. Int J Cancer 1969;4:350–63.
24. Tretli S. Height and weight in relation to breast cancer morbidity and mortality. A prospective study of 570,000 women in Norway. Int J Cancer 1989;44:23–30.
25. Ursin G, Longnecker MP, Haile RW, Greenland S. A meta analysis of body mass index and risk of premenopausal breast cancer. Epidemiology 1995;6:137–41.
26. Hunter DJ, Willett WC. Diet, body size, and breast cancer. Epidemiol Rev 1993;15:110–32.
27. Magnusson C, Baron J, Persson I, Wolk A, Bergstrom R, Trichopoulos D, Adami HO. Body size in different periods of life and breast cancer risk in postmenopausal women. Int J Cancer 1998;76:29–34.
28. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, Jong RA, Hislop G, Chiarelli A, Minkin S, Yaffe MJ. Mammographic density and the risk and detection of breast cancer. New Engl J Med 2007;356:227–36.
29. Dupont WD, Parl FF, Hartmann WH, Brinton LA, Winfield AC, Worrell JA, Schuyler PA, Plummer WD. Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. Cancer 1993;71:1258–65.
30. Singletary SE. Rating the risk factors for breast cancer. Ann Surg 2003;237:474–82.
31. Colditz GA, Willett WC, Hunter DJ, Stampfer MJ, Manson JE, Hennekens CH, Rosner BA. Family history, age, and risk of breast cancer. Prospective data from the Nurses' Health Study. JAMA 1993;270:338–43.
32. Claus EB, Schildkraut JM, Thompson WD, Risch NJ. The genetic attributable risk of breast and ovarian cancer. Cancer 1996;77:2318–24.
33. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, et al. A genome wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007;39:870–4.
34. Hartman M, Czene K, Reilly M, Bergh J, Lagiou P, Trichopoulos D, Adami HO, Hall P. Genetic implications of bilateral breast cancer: a population based cohort study. Lancet Oncol 2005;6:377–82.
35. Monninkhof EM, Elias SG, Vlems FA, van der Tweel I, Schuit AJ, Voskuil DW, van Leeuwen FE; TFPAC. Physical activity and breast cancer: a systematic review. Epidemiology 2007;18:137–57.
36. World Cancer Research Fund and American Institute for Cancer Prevention. Food, Nutrition and the Prevention of Cancer: a global perspective. Washington DC: World Cancer Research Fund and American Institute for Cancer Prevention, 1997;252.
37. Trichopoulou A. Olive oil and breast cancer. Cancer Causes Control 1995;6:475–6.
38. Dorgan JF, Reichman ME, Judd JT, Brown C, Longcope C, Schatzkin A, Forman M, Campbell WS, Franz C, Kahle L, Taylor PR. Relation of energy, fat, and fiber intakes to plasma concentrations of estrogens and androgens in premenopausal women. Am J Clin Nutr 1996;64:25–31.
39. Carruba G, Granata OM, Pala V, Campisi I, Agostara B, Cusimano R, Ravazzolo B, Traina A. A traditional Mediterranean diet decreases endogenous estrogens in healthy postmenopausal women. Nutr Cancer 2006;56:253–9.
40. Thiebaut AC, Kipnis V, Chang SC, Subar AF, Thompson FE, Rosenberg PS, Hollenbeck AR, Leitzmann M, Schatzkin A. Dietary fat and postmenopausal invasive breast cancer in the National Institutes of Health AARP Diet and Health Study cohort. J Natl Cancer Inst 2007;99:451–62.
41. Singletary KW, Gapstur SM. Alcohol and breast cancer: review of epidemiologic and experimental evidence and potential mechanisms. JAMA 2001;286:2143–51.
42. Trichopoulos D, Lagiou A. Tobacco smoking and cancer of the breast. In: Boyle P, Gray N, Henningfield J, Seffrin J, Zatonski W, eds. Tobacco and public health: science and policy. Oxford: Oxford University Press, 2004;503–510.
43. Gram IT, Braaten T, Terry PD, Sasco AJ, Adami HO, Lund E, Weiderpass E. Breast cancer risk among women who start smoking as teenagers. Cancer Epidemiol Biomarkers Prev 2005;14:61–6.
44. Laden F, Hankinson SE, Wolff MS, Colditz GA, Willett WC, Speizer FE, Hunter DJ. Plasma organochlorine levels and the risk of breast cancer: an extended follow up in the Nurses' Health Study. Int J Cancer 2001;91:568–74.
45. Laden F, Neas LM, Tolbert PE, Holmes MD, Hankinson SE, Spiegelman D, Speizer FE, Hunter DJ. Electric blanket use and breast cancer in the Nurses' Health Study. Am J Epidemiol 2000;152:41–9.
46. Miyoshi Y, Funahashi T, Kihara S, Taguchi T, Tamaki Y, Matsuzawa Y, Noguchi S. Association of serum adiponectin levels with breast cancer risk. Clin Cancer Res 2003;9:5699–704.
47. Tworoger SS, Eliassen AH, Kelesidis T, Colditz GA, Willett WC, Mantzoros CS, Hankinson SE. Plasma adiponectin concentrations and risk of incident breast cancer. J Clin Endocrinol Metab 2007;92:1510–16.
48. Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst 2002;94:606–16.
49. Kaaks R, Rinaldi S, Key TJ, Berrino F, Peeters PH, Biessy C, Dossus L, Lukanova A, Bingham S, Khaw KT, Allen NE, Bueno de Mesquita HB, et al. Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. Endocr Relat Cancer 2005;12:1071–82.
50. Adami HO, Johansson ED, Vegelius J, Victor A. Serum concentrations of estrone, androstenedione, testosterone and sex hormone binding globulin in postmenopausal women with breast cancer and in age matched controls. Ups J Med Sci 1979;84:259–74.
51. Kaaks R, Berrino F, Key T, Rinaldi S, Dossus L, Biessy C, Secreto G, Amiano P, Bingham S, Boeing H, Bueno de Mesquita HB, Chang Claude J, et al. Serum sex steroids in premenopausal women and breast cancer risk within the European Prospective Investigation into Cancer and Nutrition (EPIC). J Natl Cancer Inst 2005;97:755–65.
52. Renehan AG, Zwahlen M, Minder C, O'Dwyer ST, Shalet SM, Egger M. Insulin like growth factor (IGF) I. IGF binding protein 3, and cancer risk: systematic review and meta regression analysis. Lancet 2004;363:1346–53.
53. Trichopoulos D. Hypothesis: does breast cancer originate in utero? Lancet 1990;335:939–40.
54. Ekbom A, Trichopoulos D, Adami H O, Hsieh C c, Lan S j. Evidence of prenatal influences on breast cancer risk. Lancet 1992;340:1015–18.

55. Adami H O, Persson I, Ekbom A, Wolk A, Pontén J, Trichopoulos D. The aetiology and pathogenesis of human breast cancer. Mutat Res 1995;333:29 35.
56. Trichopoulos D, Lipworth L. Editorial: is cancer causation simpler than we thought, but more intractable? Epidemiology 1995;6: 347 9.
57. Lagiou P, Hsieh CC, Trichopoulos D, Xu B, Wuu J, Mucci L, Tamimi R, Adami HO, Cnattingius S. Birthweight differences between USA and China and their relevance to breast cancer aetiology. Int J Epide miol 2003;32:193 8.
58. Trichopoulos D. Intrauterine environment, mammary gland mass and breast cancer risk. Breast Cancer Res 2003;5:42 4.
59. Cole P, MacMahon B. Oestrogen fractions during early reproductive life in the aetiology of breast cancer. Lancet 1969;1:604 6.
60. DeWaard F, Trichopoulos D. A unifying concept of the aetiology of breast cancer. Int J Cancer 1988;41:666 9.
61. Le Marchand L, Kolonel LN, Myers BC, Mi MP. Birth characteristics of premenopausal women with breast cancer. Br J Cancer 1988;57: 437 9.
62. Russo J, Russo IH. Development of the human mammary gland. In: Neville MC, Daniel CW, eds. The mammary gland. New York: Ple num Press, 1987.67 93.
63. Trichopoulos D, Lipman R. Mammary gland mass and breast cancer risk. Epidemiology 1992;3:523 6.
64. Baik I, Becker PS, DeVito WJ, Lagiou P, Ballen K, Quesenberry PJ, Hsieh CC. Stem cells and prenatal origin of breast cancer. Cancer Causes Control 2004;15:517 30.
65. Ekbom A, Adami H O, Trichopoulos D, Lambe M, Hsieh C c, Pontén J. Epidemiologic correlates of breast cancer laterality (Sweden). Can cer Causes Control 1994;5:510 16.
66. Trichopoulos D, Bamia C, Lagiou P, Trichopoulou A. Are mammo tropic hormones mainly permissive for the development of breast can cer? Int J Cancer 2006;118:2863 5.
67. Russo J, Romero AI, Russo IH. Architectural pattern of the normal and cancerous breast under the influence of parity. Cancer Epidemiol Biomark Prev 1994;3:219 24.
68. Willett W. Nutritional epidemiology, 2nd edn. New York: Oxford University Press, 1998.377 413.
69. Lipworth L, Hsieh C c, Wide L, Ekbom A, Yu S z, Yu G p, Xu B, Hellerstein S, Carlstrom K, Trichopoulos D, Adami H O. Maternal pregnancy hormone levels in an area with a high incidence (Boston, USA) and in an area with a low incidence (Shanghai, China) of breast cancer. Brit J Cancer 1999;79:7 12.
70. Lagiou P, Samoli E, Lagiou A, Hsieh CC, Adami HO, Trichopoulos D. Maternal height, pregnancy estriol and birth weight in reference to breast cancer risk in Boston and Shanghai. Int J Cancer 2005;117: 494 8.
71. Baik I, Devito WJ, Ballen K, Becker PS, Okulicz W, Liu Q, Delpapa E, Lagiou P, Sturgeon S, Trichopoulos D, Quesenberry PJ, Hsieh CC. Association of fetal hormone levels with stem cell potential: evidence for early life roots of human cancer. Cancer Res 2005;65:358 63.
72. Savarese TM, Strohsnitter WC, Low HP, Liu Q, Baik I, Okulicz W, Chelmow DP, Lagiou P, Quesenberry PJ, Noller KL, Hsieh CC Corre lation of umbilical cord blood hormones and growth factors with stem cell potential: implications for the prenatal origin of breast cancer hypothesis. Breast Cancer Res 2007;9:R29.
73. De Assis S, Hilakivi Clarke L. Timing of dietary estrogenic exposures and breast cancer risk. Ann NY Acad Sci 2006;1089:14 35.
74. Hilakivi Clarke L, de Assis S. Fetal origins of breast cancer. Trends Endocrinol Metab 2006;17:340 8.
75. Palmer JR, Hatch EE, Rosenberg CL, Hartge P, Kaufman RH, Titus Ernstoff L, Noller KL, Herbst AL, Rao RS, Troisi R, Colton T, Hoo ver RN. Risk of breast cancer in women exposed to diethylstilbestrol in utero: prelimiinary results (United States). Cancer Causes Control 2002;13:753 8.
76. Palmer JR, Wise LA, Hatch EE, Troisi R, Titus Ernstoff L, Strohsnit ter W, Kaufman R, Herbst AL, Noller KL, Hyer M, Hoover RN. Pre natal diethylstilbestrol exposure and risk of breast cancer. Cancer Epi demiol Biomarkers Prev 2006;15:1509 14.
77. Ekbom A, Thurfjell E, Hsieh CC, Trichopoulos D, Adami HO. Peri natal characteristics and adult mammographic patterns. Int J Cancer 1995;61:177 80.
78. Cerhan JR, Sellers TA, Janney CA, Pankratz VS, Brandt KR, Vachon CM. Prenatal and perinatal correlates of adult mammographic breast density. Cancer Epidemiol Biomarkers Prev 2005;14:1502 8.

OPEN ACCESS Freely available online

PLoS MEDICINE

Perspective

Birth Size and the Pathogenesis of Breast Cancer

Pagona Lagiou, Dimitrios Trichopoulos*

In a major undertaking reported in this issue of *PLoS Medicine*, a collaborative group co-ordinated by Isobel dos Santos Silva provide all but conclusive evidence that birth size is a predictor of breast cancer risk in adult life [1]. The researchers compiled and reanalysed individual participant data from 32 studies, comprising 22,058 cases of breast cancer. On the basis of reliable data retrieved from birth records, they found that an increase of birth weight by 500 grams was associated with a statistically significant 6% increase in breast cancer risk; whereas, controlling for birth weight, an increase of birth length by two centimetres was associated with a 9% increase in this risk. The relative size of the effects is small, but the individual studies driving the conclusion were of sound epidemiological design (cohort or nested case-control) and relied on objectively documented birth size parameters, allowing little room for selection or information bias. Now that the question of whether birth size is associated with breast cancer risk appears to be settled, a number of additional questions need to be addressed.

How Important Are These Findings in Biological and Practical Terms?

In practical terms, a 10% increase in breast cancer risk at the higher birth size category is certainly small, but not trivial for a common disease like breast cancer. Indeed, the gradient is in the same order of magnitude as that found for other common risk factors for breast cancer, such as age at menarche, age at menopause, or postmenopausal obesity [2]. And, from a biological point of view, it is certainly important to document a phenomenon that indicates the involvement of intrauterine processes in a major human cancer, as has already been done in animal models [3].

The Perspective section is for experts to discuss the clinical practice or public health implications of a published article that is freely available online.

Linked Research Article

This Perspective discusses the following new study published in *PLoS Medicine*:

dos Santos Silva I, De Stavola B, McCormack V, Collaborative Group on Pre-Natal Risk Factors and Subsequent Risk of Breast Cancer (2008) Birth size and breast cancer risk: Re-analysis of individual participant data from 32 studies. PLoS Med 5(9): e193. doi:10.1371/journal.pmed.0050193

Isobel dos Santos Silva and Bianca De Stavola and colleagues reanalyzed individual participant data from 32 published and unpublished studies to obtain precise estimates of the association between birth size and breast cancer risk.

Can the Results Help To Explain Patterns in Breast Cancer Incidence Around the World?

The observation of sharp ecological contrasts in breast cancer incidence around the world is one of the most challenging features in the epidemiology of the disease, and every hypothesis on the aetiology of breast cancer should be able to accommodate these contrasts. Breast cancer incidence among women of European descent in the Western world is several times higher than that among Chinese or Japanese women in Asia. The gradual elimination of this difference over several generations among Asian migrants in Western countries implies that genetic factors are not responsible for the ecological contrasts [2].

In our view, the results of the new collaborative group study [1] are compatible with the ecological patterns of breast cancer incidence. Newborns in China have lower birth weight than newborns of European descent in the United States, largely due to differences in maternal anthropometry that impose physical constraints on newborn size [4]. Migration from China to the US is associated with increased energy intake, leading to increased adult body size (including pelvic size), and consequently to the removal of constraints on birth weight. The cycle tends to repeat itself over consecutive generations of Asians migrating to the West and is associated with a gradual increase of breast cancer incidence in this population [4]. In the collaborative group study [1], controlling for adult height only slightly reduced the association of birth size with breast cancer risk, but, as the authors indicate, the adjustment was based on a small number of cases and misclassification may have hindered documentation of an important mediating role of adult height.

Funding: The authors are supported by the W81XWH-05-1-0314 Innovator Award from the US Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs. The funder played no role in the preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Citation: Lagiou P, Trichopoulos D (2008) Birth size and the pathogenesis of breast cancer. PLoS Med 5(9): e194. doi:10.1371/journal.pmed.0050194

Pagona Lagiou and Dimitrios Trichopoulos are in the Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America. Pagona Lagiou is also in the Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, University of Athens, Athens, Greece.

* To whom correspondence should be addressed. E-mail: dtrichop@hsph.harvard.edu

Provenance: Commissioned; not externally peer reviewed

Can We Integrate the Study's Findings into Our Current Understanding of the Early Stages in the Natural History of Breast Cancer?

Our current understanding of the early stages in the natural history of breast cancer is limited. Hilakivi-Clarke and de Assis have suggested that epigenetic modifications associated with large birth size lead to modifications in mammary gland development and increased vulnerability of epithelial targets for malignant transformation [3]. It has also been postulated that higher birth size is associated with higher levels of pregnancy hormones, including estrogens and insulin-like growth factor 1, which favour the generation of a higher number of susceptible stem cells with compromised genomic stability [5]. In this context, mammary gland mass, an important determinant of breast cancer risk, could be viewed as an adult life correlate of the number of mammary cells susceptible to transformation [5,6]. The group led by Chung-Cheng Hsieh of the University of Massachusetts is doing important work in this field. This group has reported that high levels of insulin-like growth factor 1 and estriol are associated with larger pools of stem cells in the cord blood [7], and that birth size is also associated with the stem cell pool [8].

Are There Implications for the Primary Prevention of Breast Cancer?

Documentation of a positive association of birth size, particularly birth length, with breast cancer risk in adult life may improve prediction of disease risk, but does not offer much opportunity for prevention, particularly since birth size is inversely associated with cardiovascular risk [9]. The situation could change if other periods in early life, particularly postnatal life, were found to be related to adult life breast cancer risk. In any case, recognition of early life influences as critical in the aetiology of breast cancer helps to explain why several adult life primary prevention practices have been of limited effectiveness.

Are Perinatal Exposures Important for Breast Cancer Only, Or Could They Affect Risk of Other Cancers As Well?

The mammary gland seems to be the only organ that is not fully developed at birth [10], which implies that mammary tissue–specific stem cells may remain in a quiescent stage for longer periods than tissue-specific stem cells for other organs. This could provide an explanation for why intrauterine factors are more important for breast cancer than for other cancers. It is reasonable, however, to expect that intrauterine factors could affect the risk of other forms of cancer, albeit to a lesser extent. In fact, weak birth weight associations have been reported for other cancers, although the evidence is still limited [11].

The intrauterine life has been implicated in the aetiology of breast cancer on the basis of theoretical arguments and epidemiological considerations [12]. However, the documentation of its role in breast cancer risk has relied on studies linking birth size to this risk. The Collaborative Group on Pre-Natal Risk Factors and Subsequent Risk of Breast Cancer has elegantly and most efficiently reanalysed these studies and, by pooling together data at the individual level, has provided the strongest evidence yet that birth size is a critical determinant of breast cancer risk in adult life. ■

References

1. dos Santos Silva I, De Stavola B, McCormack V, Collaborative Group on Pre-Natal Risk Factors and Subsequent Risk of Breast Cancer (2008) Birth size and breast cancer risk: Re-analysis of individual participant data from 32 studies. PLoS Med 5: e193. doi:10.1371/journal.pmed.0050193
2. Hankinson S, Tamimi R, Hunter D (2008) Breast cancer. In: Adami HO, Hunter D, Trichopoulos D, editors. Textbook of cancer epidemiology. 2nd edition. Oxford: Oxford University Press. pp 403-445.
3. Hilakivi-Clarke L, de Assis S (2006) Fetal origins of breast cancer. Trends Endocrinol Metab 17: 340-348.
4. Lagiou P, Hsieh CC, Trichopoulos D, Xu B, Wuu J, et al. (2003) Birthweight differences between USA and China and their relevance to breast cancer aetiology. Int J Epidemiol 32: 193-198.
5. Trichopoulos D, Lipman R (1992) Mammary gland mass and breast cancer risk. Epidemiology 3: 523-526.
6. Adami H-O, Persson I, Ekbom A, Wolk A, Pontén J, et al. (1995) The aetiology and pathogenesis of human breast cancer. Mutat Res 333: 29-35.
7. Savarese TM, Strohsnitter WC, Low HP, Liu Q, Baik I, et al. (2007) Correlation of umbilical cord blood hormones and growth factors with stem cell potential: Implications for the prenatal origin of breast cancer hypothesis. Breast Cancer Res 9: R29.
8. Strohsnitter WC, Savarese TM, Low HP, Chelmow DP, Lagiou P, et al. (2008) Correlation of umbilical cord blood haematopoietic stem and progenitor cell levels with birth weight: Implications for a prenatal influence on cancer risk. Br J Cancer 98: 660-663.
9. Barker DJ, Winter PD, Osmond C, Margetts B, Simmonds SJ (1989) Weight in infancy and death from ischaemic heart disease. Lancet 2: 577-580.
10. Russo J, Rivera R, Russo IH (1992) Influence of age and parity on the development of the human breast. Breast Cancer Res Treat 23: 211-218.
11. Ahlgren M, Wohlfahrt J, Olsen LW, Sørensen TI, Melbye M (2007) Birth weight and risk of cancer. Cancer 110: 412-419.
12. Trichopoulos D (1990) Hypothesis: Does breast cancer originate in utero? Lancet 335: 939-940.

APPENDIX 4

PDF copies of publications on indirectly relevant research questions
with reference to support by the W81XWH-05-1-0314 Innovator Award

Hypothesis/Commentary

Plasma Volume Expansion in Pregnancy: Implications for Biomarkers in Population Studies

Jessica M. Faupel-Badger,[1,2] Chung-Cheng Hsieh,[5] Rebecca Troisi,[3] Pagona Lagiou,[6,7] and Nancy Potischman[4]

[1]Division of Cancer Prevention, [2]Mammary Biology and Tumorigenesis Laboratory, Center for Cancer Research, [3]Division of Cancer Epidemiology and Genetics, and [4]Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland; [5]University of Massachusetts Medical School, Worcester, Massachusetts; [6]Department of Hygiene and Epidemiology, University of Athens Medical School, Athens, Greece; and [7]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts

Abstract

There is a growing body of literature focused on endogenous hormone exposures during pregnancy and subsequent cancer risk for both mother and offspring. Examples of these studies include those focused on the biological mechanism for the association of preeclampsia with reduced risk of breast cancer for mother and female offspring or studies that have examined hormone concentrations during pregnancy between different ethnic groups who vary in their rates of breast cancer incidence. Although these studies seem relatively straightforward in conception and analysis, measurement of the concentration of hormones and other biomarkers in pregnant subjects is influenced by plasma volume expansion (PVE). During pregnancy, the maternal plasma volume expands 45% on average to provide for the greater circulatory needs of the maternal organs. Consequently, serum protein and hormone concentrations are greatly altered when comparing the pregnant with nonpregnant state. Assessing PVE also is complicated by the vast individual variation in PVE, ranging from minimal to a 2-fold increase. We propose that PVE needs to be evaluated when comparing biomarker concentrations during pregnancy in two populations that may differ with respect to PVE. Small body size is associated with lower PVE compared with higher body size. Therefore, we hypothesize that variation in PVE will influence the interpretation of differences in biomarker concentrations across population groups with respect to the etiologic significance of the biomarker to the disease under study (e.g., breast cancer). It is possible that some observations may be due only to differences in dilution between the two groups. We present PVE as a topic for consideration in population-based studies, examples of the types of studies where PVE may be relevant, and our own analysis of one such study in the text below. (Cancer Epidemiol Biomarkers Prev 2007;16(9):1720 3)

Introduction to Plasma Volume Expansion in Pregnancy

During pregnancy, maternal plasma volume increases to meet the greater circulatory needs of the placenta and maternal organs (e.g., uterus, breasts, skin, and kidneys), with an average increase of ~45% (1-5). There are vast differences among women, however, from a minimal change to a doubling in plasma volume (1, 6, 7). Several factors can influence plasma volume expansion (PVE) including maternal pre-pregnancy body mass index (BMI; ref. 8). In studies that have examined ethnic differences in PVE, populations that are on average shorter and weigh less exhibit less absolute change in plasma volume (9-11). This was well documented in a report comparing plasma volumes between Indians and Europeans (10). Furthermore, of studies examining PVE across different ethnicities, all show vast individual variation in PVE as well (9-11). Much of the variation is unexplained, although it is known that PVE is positively associated with parity (12), multiple births (13-15), higher birth weight (16, 17), and increased maternal pre-pregnancy BMI (8) and inversely associated with conditions of decreased fetal growth (e.g., intrauterine growth restriction) and compromised placental development (e.g., preeclampsia; refs. 1, 2, 18-20).

In clinical practice, PVE or 'hemodilution' is addressed by modifying the criteria for diagnostic biomarkers. For example, the cut-point for anemia or iron deficiency, which is defined by hemoglobin <12.0 g/dL in the

Received 4/3/07; revised 5/29/07; accepted 6/22/07.

Grant support: Division of Cancer Prevention and the Center for Cancer Research, National Cancer Institute (J.M. Faupel-Badger) and W81XWH-05-1-0314 Innovator Award and U.S. Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs (P. Lagiou and C-C. Hsieh). J.M. Faupel-Badger is a fellow in the National Cancer Institute Cancer Prevention Fellowship Program.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Jessica M. Faupel-Badger, Mammary Biology and Tumorigenesis Laboratory, Center for Cancer Research, National Cancer Institute, MSC 4254, 37 Convent Drive, Bethesda, MD 20892. Phone: 301-443-3076; Fax: 301-402-0711. E-mail: badgerje@mail.nih.gov

doi:10.1158/1055-9965.EPI-07-0311

nonpregnant state, is changed to <11.0 g/dL in the first and third trimesters and <10.5 g/dL in the second trimester. This corresponds to the increase in plasma volume starting at 6 to 10 weeks of gestation that rises sharply through the second trimester, before beginning to plateau at 32 weeks (1).

PVE and Population Studies

Highlighting diagnosis of anemia in clinical practice shows that PVE can have significant effects on biomarker concentrations. The implications of interindividual and between-group variability in PVE, however, have generally not been addressed in large population-based research studies involving pregnant subjects. In etiologic studies, this variability could introduce bias in the form of confounding if it were related to both the factor under study (e.g., preeclampsia) and the biomarker. Thus, accounting for PVE would be essential for a study focused on elucidating biomarkers involved in the causal pathway of a disease. For example, in preeclampsia, there is less PVE than in uncomplicated pregnancies (1, 2, 19). High concentrations of a biomarker such as sex hormone binding globulin (21) or soluble fms-like tyrosine kinase (1, 22, 23) among preeclamptic women compared with women who have uncomplicated pregnancies may reflect an etiologically relevant difference and/or the lower hemodilution in preeclampsia. We hypothesize that variation in PVE will affect the interpretation of differences in biomarker concentrations between individuals or population groups, especially with respect to etiologic significance, and that evaluation of PVE should be considered in population studies involving pregnant women.

Other examples in which individual variation in PVE may affect interpretation include when biomarkers are used to measure the success of nutritional intervention during pregnancy (24). In some areas of Asia, it is not uncommon for pregnant women to be deficient in multiple micronutrients (24, 25). To address this issue, one study among rural pregnant women in Nepal evaluated the effects of micronutrient supplementation on serum retinol, folate, riboflavin, and 25-hydroxyvitamin D concentrations (26). Assuming this population has wide variation in individual PVE, as has been shown in other populations, PVE may affect the perceived success of this intervention in individual women. Furthermore, variability in nutritional status and hydration across individuals could influence the apparent concentrations of nutrients of interest in this study via effects on PVE.

Measurement of PVE, Biomarker Concentrations, and Relation to Outcomes

PVE also may be an important factor in studies focused on understanding what biological features mediate associations of pregnancy characteristics with subsequent development of chronic disease in either the mother or offspring. For example, preeclampsia is associated with decreased breast cancer risk for both the mother and female offspring (27). Population-based studies of circulating biomarkers in preeclamptic and normotensive pregnancies have been conducted to elucidate possible biological mechanisms. Serum androgen concentrations at delivery are observed to be higher in preeclamptic women than in those with normal pregnancies (28). Progesterone concentrations also seem higher in maternal serum during the 27th week of pregnancy when comparing preeclamptic with normal pregnancies (29). Could the lower PVE that preeclamptic women experience relative to normotensive women explain their apparently higher concentration of serum hormones? And if so, are differences due to PVE of etiologic importance if the target tissues are exposed to the same total amount of the hormones?

Direct methods of measuring plasma volume are labor intensive and not easily adapted to large studies. Methods to approximate plasma volume involve labeling of albumin, with Evan's blue dye being the most common technique (11, 30). This method requires sampling plasma after injecting the patient with Evan's blue dye and allowing time (a minimum of 10 min) for sufficient mixing of the dye with plasma. The decay of Evan's blue dye is measured by the absorbance at 610 nm and plasma volume in milliliters per kilogram can be calculated from the absorbance measurement (30). In addition to patient monitoring during the dye injection and sample collection, it also is recommended that the patient observe a 30-min rest period before injection and that women in late pregnancy lay on their sides during injection to promote mixing of the dye (11, 30). Measuring Evan's blue dye at one point in time can approximate a measurement of plasma volume; however, to measure the increase in plasma volume, an individual would need to have the Evan's blue dye measurement completed more than once during pregnancy. These measurements across time points would then be used to determine the amount of PVE for the individual. This technique involves significant patient burden and is not practical for large studies.

In population-based studies, accounting for PVE in the data analysis by adjustment for factors highly predictive of PVE is more feasible. We reanalyzed data from a published study comparing maternal hormone concentrations in women from Boston and Shanghai (31) to determine if adjusting for correlates of PVE would affect the observed differences in estradiol concentrations. The women for this study were recruited from urban clinics affiliated with Beth Israel Hospital in Boston and three urban and one rural clinic in China affiliated with Shanghai Medical University. The women were all under the age of 40 years and were either Caucasian in Boston or Chinese in Shanghai. The purpose of the published study was to explore the hypothesis that fetal exposures, such as high *in utero* hormone concentrations, may be associated with the development of breast cancer in the offspring, using two populations with different disease incidence (31). Two serum samples, collected at 16 and 27 weeks of gestation, were used to evaluate estradiol, estriol, prolactin, progesterone, growth hormone, albumin, and sex hormone binding globulin. Given that women in Shanghai have approximately one fifth the incidence of breast cancer of women in Boston (32), it was hypothesized that the women from Shanghai would have lower hormone concentrations. In contrast, for every compound measured, the women from Shanghai had significantly higher concentrations with the exception of progesterone at 27 weeks. The results from this

Table 1. Percentage difference in maternal estradiol at 16 wks of gestation comparing Chinese with U.S. concentrations by strata of pre-pregnancy BMI, uniparous subjects only

	Adjusted for:			With further adjustment for:				
Pre pregnancy BMI (kg/m^2)	Boston (*n*)	Shanghai (*n*)	Maternal age, gestational age	Albumin	Albumin, pre weight	Albumin, height	Albumin, pre BMI	Albumin, birth weight
<19.1	21	84	37.3*	39.5*	38.2*	35.2*	40.5*	40.8*
19.1 20.5	30	68	24.5*	31.0*	19.8	20.5	30.8*	31.0*
20.6 22.2	34	46	12.6	16.4	15.9	16.6	16.7	14.7
22.3+	36	22	34.8*	35.6*	34.3*	37.7*	34.3*	35.8*
All subjects	121	200	26.7*	30.5*	27.7*	29.0*	29.4*	30.6*

NOTE: Adjusted for maternal age, pre-pregnancy weight, pre-pregnancy BMI, height, and albumin measured at week 16, and offspring birthweight and gestational age.
*$P < 0.05$.

study were unexpected as Chinese women are known to have lower circulating estradiol concentrations than Caucasian women in the nonpregnant state (33-35).

We hypothesized that differences in PVE between the Chinese and U.S. populations could explain the higher estradiol concentrations observed among the Chinese. Asian women have lower values for several correlates of PVE, such as height, weight, and infant birth weight, when compared with Caucasian women (31, 36, 37). To assess whether the higher hormone concentrations present in Asian women were a result of less PVE compared with Caucasian women, we analyzed the data set from Lipworth et al. (31) adjusting for or stratifying on correlates of PVE. The analysis was restricted to uniparous (i.e., first pregnancy) women because parity is known to affect PVE (11). Regression analysis with log-transformed estradiol as the dependent variable and with adjustment for maternal age, pre-pregnancy weight, height, pre-pregnancy BMI, birth weight, and albumin concentration allowed calculation of percentage difference in estradiol concentrations between the Shanghai and Boston populations for subjects with complete data on covariates (131 subjects for Boston and 220 in Shanghai). Similar to results in the original study, among all uniparous women in our analysis, the estradiol concentrations in the women from Shanghai were 26.7% higher than the estradiol concentrations in the women from Boston, at 16 weeks of gestation, after adjusting for maternal age and gestational week. Table 1 presents the percentage difference among women at 16 weeks of gestation within the same category of pre-pregnancy BMI to attempt to further account for PVE. The differences in estradiol concentrations between the Boston and Shanghai populations were strongest at 16 weeks and adjustment for correlates of PVE, including maternal pre-pregnancy BMI and birth weight, had no effect. Results for analyses of estradiol concentrations at week 16 were similar when all subjects from the Lipworth study, uniparous and multiparous, were included in the analyses (data not shown).

Conclusions and Comments for Future Analysis of PVE

Our reanalysis of the estradiol data shows that the difference in estradiol concentrations between the Boston and Shanghai populations cannot be explained by adjusting for currently accepted correlates of PVE. Therefore, we have two possible conclusions. The difference in estradiol concentrations between these two populations may be real and due to factors other than differences in PVE. Alternatively, it is possible that pre-pregnancy BMI may not be a sufficient marker of PVE. In nonpregnant women, it may be possible to adjust for differences in plasma volume by including BMI in the model, as BMI is highly correlated with plasma volume in healthy, nonpregnant individuals (8). Whereas pre-pregnancy BMI is currently an accepted correlate for PVE, other factors (e.g., physiologic processes related to adaptation to pregnancy and size of fetus) besides maternal size can affect PVE, thus suggesting that pre-pregnancy BMI may not be the most adequate correlate of PVE, although it may be the best one in use at this time.

Although reanalysis of the data from the study by Lipworth et al. did not support our hypothesis that pre-pregnancy BMI would be a sufficient predictor to account for PVE, it may be worth considering this type of analysis in other studies with larger numbers or among different ethnicities. In addition, to fully account for any effects of PVE, it may be that new proxies or methods to evaluate PVE need to be developed for practical use in large studies. While new proxies or methods are being explored, other avenues to measure hormone exposure during pregnancy could be used to circumvent the issue of PVE. For example, studies focused on *in utero* hormone exposure and subsequent disease risk in the offspring could also examine cord blood samples for the measurement of hormones or other biomarkers (38). These samples would be insulated from variations in dilution present in the maternal circulation and may be more relevant to fetal exposure than conclusions based on maternal serum. In addition to cord blood, saliva or urine may also provide a source for hormone measurements that may be less influenced by PVE. Future studies with samples to measure the correlation of maternal serum hormone concentrations to concentrations in cord blood, urine and saliva would be helpful to determine if the magnitude increase in hormone concentrations seen in serum is reflected in the alternative method. It would also be necessary to evaluate whether the alternative method may also be influenced by PVE. Another approach would be to evaluate changes in a biomarker over the course of a pregnancy, perhaps elucidating different patterns in those that develop a pregnancy condition from those who remain normal.

In conclusion, given the vast individual and group differences in PVE, its implications for the interpretation of results from biomarkers studies should be considered. We hope this commentary begins a dialogue about the implications of PVE and its evaluation for the types of studies outlined above. Further evaluation of PVE is necessary to determine its potential effect on the interpretation of observed results with regard to etiology and the degree to which it needs to be addressed in study design and/or analysis.

References

1. Blackburn ST, Loper DL. Maternal, fetal, and neonatal physiology. A clinical perspective. Philadelphia: W.B. Saunders Co.; 1992. p. 160 2, 171, 202 3, 222 8.
2. Brinkman CR. Maternal cardiovascular and renal disorders. Maternal-fetal medicine. Philadelphia: W.B. Saunders; 1984. p. 679 90.
3. Chesley LC. Plasma and red cell volumes during pregnancy. Am J Obstet Gynecol 1972;112:440.
4. Peck TM, Arais F. Hematologic changes associated with pregnancy. Clin Obstet Gynecol 1979;22:785.
5. Pritchard JA. Changes in blood volume during pregnancy and delivery. Anesthesiology 1965;26:393.
6. Pritchard JA, MacDonald PC, Gant N. Williams' obstetrics. New York: Appleton-Century-Crofts; 1985. p. 191 4.
7. Anderson HM. Maternal hematologic disorders. Maternal-fetal medicine. Philadelphia: W.B. Saunders; 1984. p. 795 832.
8. Dagher FJ, Lyons JH, Finlayson DC, Shamsai J, Moore FD. Blood volume measurement: a critical study prediction of normal values: controlled measurement of sequential changes: choice of a bedside method. Adv Surg 1965;1:69 109.
9. Harrison KA. Blood volume changes in normal pregnant Nigerian women. J Obstet Gynaecol Br Commonw 1966;73:717 23.
10. Hutchins CJ. Plasma volume changes in pregnancy in Indian and European primigravidae. Br J Obstet Gynaecol 1980;87:586 9.
11. Hytten F. Blood volume changes in normal pregnancy. Clin Haematol 1985;14:601 12.
12. Campbell DM, MacGillivray I. Comparison of maternal response in first and second pregnancies in relation to baby weight. J Obstet Gynaecol Br Commonw 1972;79:684 93.
13. Fullerton WT, Hytten FE, Klopper AI, McKay E. A case of quadruplet pregnancy. J Obstet Gynaecol Br Commonw 1965;72:791 6.
14. Letsky E. The haematological system. In: Hytten FE, Chamberlain GVP, editors. Clinical physiology in obstetrics. Oxford: Blackwell Scientific Publications; 1980. p. 43 78.
15. Rovinsky JJ, Jaffin H. Cardiovascular hemodynamics in pregnancy. I. Blood and plasma volumes in multiple pregnancy. Am J Obstet Gynecol 1965;93:1 15.
16. Hytten FE, Paintin DB. Increase in plasma volume during normal pregnancy. J Obstet Gynaecol Br Emp 1963;70:402 7.
17. Pirani BB, Campbell DM, MacGillivray I. Plasma volume in normal first pregnancy. J Obstet Gynaecol Br Commonw 1973;80: 884 7.
18. Bentley DP. Iron metabolism and anaemia in pregnancy. Clin Haematol 1985;14:613.
19. Salas SP, Marshall G, Gutierrez BL, Rosso P. Time course of maternal plasma volume and hormonal changes in women with preeclampsia or fetal growth restriction. Hypertension 2006;47:203 8.
20. Salas SP, Rosso P, Espinoza R, et al. Maternal plasma volume expansion and hormonal changes in women with idiopathic fetal growth retardation. Obstet Gynecol 1993;81:1029 33.
21. Wolf M, Sandler L, Munoz K, et al. First trimester insulin resistance and subsequent preeclampsia: a prospective study. J Clin Endocrinol Metab 2002;87:1563 8.
22. Levine RJ, Lam C, Qian C, et al. Soluble endoglin and other circulating antiangiogenic factors in preeclampsia. N Engl J Med 2006;355:992 1005.
23. Levine RJ, Maynard SE, Qian C, et al. Circulating angiogenic factors and the risk of preeclampsia. N Engl J Med 2004;350:672 83.
24. Ahmed F, Mahmuda I, Sattar A, Akhtaruzzaman M. Anaemia and vitamin A deficiency in poor urban pregnant women of Bangladesh. Asia Pac J Clin Nutr 2003;12:460 6.
25. Jiang T, Christian P, Khatry SK, Wu L, West KP, Jr. Micronutrient deficiencies in early pregnancy are common, concurrent, and vary by season among rural Nepali pregnant women. J Nutr 2005;135:1106 12.
26. Christian P, Jiang T, Khatry SK, et al. Antenatal supplementation with micronutrients and biochemical indicators of status and subclinical infection in rural Nepal. Am J Clin Nutr 2006;83:788 94.
27. Innes KE, Byers TE. Preeclampsia and breast cancer risk. Epidemiology 1999;10:722 32.
28. Troisi R, Potischman N, Roberts JM, et al. Maternal serum oestrogen and androgen concentrations in preeclamptic and uncomplicated pregnancies. Int J Epidemiol 2003;32:455 60.
29. Tamimi R, Lagiou P, Vatten LJ, et al. Pregnancy hormones, preeclampsia, and implications for breast cancer risk in the offspring. Cancer Epidemiol Biomarkers Prev 2003;12:647 50.
30. el-Sayed H, Goodall SR, Hainsworth R. Re-evaluation of Evans blue dye dilution method of plasma volume measurement. Clin Lab Haematol 1995;17:189 94.
31. Lipworth L, Hsieh CC, Wide L, et al. Maternal pregnancy hormone levels in an area with a high incidence (Boston, USA) and in an area with a low incidence (Shanghai, China) of breast cancer. Br J Cancer 1999;79:7 12.
32. Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2002: Cancer incidence, mortality and prevalence Worldwide. In: IARC, editor. CancerBase no. 5. version 2.0 2006. Lyon: IARC Press; 2004.
33. Bernstein L, Yuan JM, Ross RK, et al. Serum hormone levels in premenopausal Chinese women in Shanghai and white women in Los Angeles: results from two breast cancer case-control studies. Cancer Causes Control 1990;1:51 8.
34. Key TJ, Chen J, Wang DY, Pike MC, Boreham J. Sex hormones in women in rural China and in Britain. Br J Cancer 1990;62:631 6.
35. Ursin G, Wilson M, Henderson BE, et al. Do urinary estrogen metabolites reflect the differences in breast cancer risk between Singapore Chinese and United States African-American and white women? Cancer Res 2001;61:3326 9.
36. Lagiou P, Hsieh CC, Trichopoulos D, et al. Birthweight differences between USA and China and their relevance to breast cancer aetiology. Int J Epidemiol 2003;32:193 8.
37. Wen SW, Kramer MS, Usher RH. Comparison of birth weight distributions between Chinese and Caucasian infants. Am J Epidemiol 1995;141:1177 87.
38. Troisi R, Potischman N, Roberts JM, et al. Correlation of serum hormone concentrations in maternal and umbilical cord samples. Cancer Epidemiol Biomarkers Prev 2003;12:452 6.

Cord Serum Estrogens, Androgens, Insulin-Like Growth Factor-I, and Insulin-Like Growth Factor Binding Protein-3 in Chinese and U.S. Caucasian Neonates

Rebecca Troisi,[1,3] Pagona Lagiou,[4,5,7] Dimitrios Trichopoulos,[5,7] Biao Xu,[8] Lucy Chie,[6] Frank Z. Stanczyk,[9] Nancy Potischman,[2] Hans-Olov Adami,[5,7] Robert N. Hoover,[1] and Chung-Cheng Hsieh[5,7,10]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Department of Health and Human Services; [2]Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland; [3]Department of Community and Family Medicine, Dartmouth Medical School, Hanover, New Hampshire; [4]Department of Hygiene and Epidemiology, University of Athens Medical School, Athens, Greece; [5]Department of Epidemiology, Harvard School of Public Health; [6]Department of Obstetrics and Gynecology, Beth Israel Deaconess Medical Center, Boston, Massachusetts; [7]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [8]Department of Epidemiology, School of Public Health, Fudan University, Shanghai, China; [9]Reproductive Endocrine Research Laboratory, University of Southern California Keck School of Medicine, Los Angeles, California; and [10]Department of Cancer Biology, University of Massachusetts Medical Center, Worcester, Massachusetts

Abstract

Markedly lower breast cancer incidence rates in Asians than Caucasians are not explained by established adult risk factors. Migration studies suggest the importance of early-life exposures, including perhaps the *in utero* period. Concentrations of steroid hormones and insulin-like growth factors (IGF) were measured in umbilical cord sera from pregnancies in Shanghai, China (n = 121) and Boston, MA (n = 111). Pregnancy characteristics were ascertained by interview and medical records. Means and percent differences in hormone concentrations comparing Chinese with Caucasians and 95% confidence intervals were estimated from linear regression models. Cord concentrations of androstenedione (91.9%), testosterone (257%), estriol (48.6%), and IGF binding protein-3 (21.1%) were significantly higher in the Chinese than U.S. samples, and cord prolactin was lower (−14.9%). Cord estradiol and IGF-I concentrations did not differ by race/ethnicity. With adjustment for gestational length, maternal age, pre-pregnancy weight, and weight gain, androstenedione (60.5%), testosterone (185%), and IGF binding protein-3 (40.4%) remained significantly higher in the Chinese, whereas the higher estriol and lower prolactin concentrations were attenuated. In addition, estradiol levels became lower in the Chinese (−29.8%) but did not reach statistical significance. Results were generally similar when restricted to first full-term pregnancies, with reduced estradiol concentrations in the Chinese reaching statistical significance after adjustment. These data are consistent with the hypothesis that elevated prenatal androgen exposure could mediate reductions in breast cancer risk. The meaning of the change in findings for estrogens after controlling for factors related to the pregnancy is unclear with regard to explaining international breast cancer differences. (Cancer Epidemiol Biomarkers Prev 2008;17(1):224–31)

Introduction

The most pronounced variation in breast cancer rates is observed internationally. Incidence in East and Southeast Asia is nearly one-fifth of that in northern and western Europe (1), but rates gradually increase among Asian migrants to western countries (2). Breast cancer incidence rates in the first generation born in the west, however, may be substantially elevated when compared with migrants who were born in Asia but lived decades in the west (2). Furthermore, differences in incidence rates by migration status are not fully explained by established menstrual and reproductive risk factors for breast cancer (3, 4). These observations are consistent with environmental factors early in life explaining at least some of the variation in breast cancer rates across populations (5-7), and with mother's environment during pregnancy influencing the *in utero* environment.

Whereas the full range of hormones involved in breast carcinogenesis is unclear, evidence indicates the importance of estrogens, androgens, and progesterone (8), all of which rise significantly during pregnancy. Trichopoulos (9) hypothesized that exposure to lower *in utero* estrogen concentrations affords protection against subsequent breast carcinogenesis. In the nonpregnant state, circulating estrogen concentrations are generally lower in premenopausal and postmenopausal Asian women compared with Caucasian women (10, 11). Although

Received 6/12/07; revised 10/24/07; accepted 11/5/07.

Grant support: W81XWH-05-1-0314 Innovator Award, U.S. Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs (P. Lagiou, D. Trichopoulos, H-O. Adami, and C-C. Hsieh).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Rebecca Troisi, Dartmouth-Hitchcock Medical Center, Room 854, 7297 Rubin Building, One Medical Center Drive, Lebanon, NH 03756. Phone: 603-653-9024. E-mail: troisir@mail.nih.gov

doi:10.1158/1055-9965.EPI-07-0536

androgens also have been generally lower in Asian women (10), a recent study showed an inverse correlation with increasing westernization in Asian migrants to the West (12). An investigation of maternal mid-pregnancy serum hormone concentrations, however, showed *higher* estradiol and estriol levels in Chinese women compared with American women (13) as well as elevations in several other compounds, including prolactin, progesterone, human growth hormone, albumin, sex hormone-binding globulin, and possibly α-fetoprotein levels (14). However, maternal hormone concentrations may or may not be representative of fetal exposure, and investigating hormone differences in the fetal circulation is warranted.

These are the first reported data on neonates born in China and in the United States to address whether cord concentrations of several estrogens and androgens, insulin-like growth factor (IGF)-I, and IGF binding protein (IGFBP)-3 differ between countries with low and high breast cancer incidence rates.

Materials and Methods

The study has been described previously (13). Pregnant women were recruited at their first prenatal visit to collaborating maternity clinics at the Beth Israel Hospital (Boston, MA) and from hospitals affiliated with the Shanghai Medical University (Shanghai, China). All U.S. women were urban residents, whereas women from Shanghai were recruited from three urban clinics and one rural clinic. The institutional review boards in Boston and Shanghai approved the study, and informed consent was obtained from all study participants.

Included were neonates of women less than 40 years old, with at most one previous still or live-born child. Only Caucasians in the United States and Chinese in Shanghai were included, and in both places, the mothers had to be proficient in the local language. Neonates born to women who had taken any hormonal medication during the index pregnancy or who had a previous diagnosis of diabetes mellitus or thyroid disease were excluded from the study, as were those neonates with a known major anomaly.

Between March 1994 and October 1995, 402 eligible women were identified at Beth Israel Hospital. Of these women, 77 (19.2%) declined to participate. An additional 9 (2.2%) women were excluded at a later date because of early spontaneous or induced pregnancy termination, 2 (0.5%) because of a twin birth, and 10 (2.5%) were lost to follow-up after the initial meeting. In Shanghai, 424 eligible women were identified between April 1994 and May 1995. Of these, 73 (17.2%) declined to participate, 2 (0.5%) were later excluded because of induced abortion, 2 (0.5%) because of a twin birth, 5 (1.2%) because of implied gestation durations of <30 or >50 weeks, and 7 (1.7%) were lost to follow-up after the initial meeting. In total, 304 and 335 pregnant women were enrolled in the study from Boston and Shanghai, respectively.

Umbilical cord blood collection started in December 1994. At delivery, the placenta was weighed, with the cord cut at the insertion site and the extra blood and clots minimized. Mixed cord blood was collected in sterile tubes without preservatives and refrigerated at 4°C for up to 24 hours until centrifugation. Samples were transported in a cooler from the rural clinic in China to a laboratory near Shanghai Medical University where they were centrifuged on the same day and the serum was aliquoted. The aliquots were stored at −20°C for about 5 to 7 days in the laboratory before being transported to Shanghai Medical University and stored at −80°C with the samples collected at the Shanghai hospitals. At the end of the study, all samples were shipped by air on dry ice to Boston where they were stored at −80°C. In total, 246 (115 from Boston and 131 from Shanghai) cord blood samples were collected.

Analytes were measured in cord serum at the Reproductive Endocrine Research Laboratory of the University of Southern California Keck School of Medicine under the direct supervision of one of us (F.Z.S.). Levels of estradiol, testosterone, and androstenedione were measured by RIA following extraction with organic solvent and purification by Celite column partition chromatography (15-17). Estriol was measured by RIA after a dual organic solvent extraction procedure (18). Prolactin, IGF-I, and IGFBP-3 were quantified by direct chemiluminescent immunoassay using the Immulite analyzer (Siemens Medical Solutions Diagnostics, Los Angeles, CA).

Because some of the samples showed signs of hemolysis (34%), measurements were prioritized with the direct assays (that is, prolactin, IGF-I, and IGFBP-3) done in samples with the least hemolysis. After this prioritization, when volume was insufficient to accommodate assays for all of the hormones, samples were randomly assigned based on groupings that optimized the number of assays that could be done. Blinded aliquots of pooled cord sera were included with the study samples and the laboratory technicians were blinded to country of origin. The coefficients of variation for the blinded replicates were 8.5% for androstenedione, 9.4% for testosterone, 10.5% for estradiol, 11.6% for estriol, 8.2% for prolactin, 3.2% for IGF-I, and 5.0% for IGFBP-3.

Information on maternal, gestational, and perinatal characteristics was obtained from the medical record and pediatric chart and from an interview with the mother. Gestational age was defined as the time since the first day of the last menstrual period.

Hormone values that were more than three interquartile ranges above the mean were excluded as outliers. The nonparametric Wilcoxon rank-sum test was applied in univariate comparisons of the maternal, gestational, and perinatal factors between the two study sites (19). Linear regression models with logarithm-transformed hormones as the dependent variable and an indicator variable representing the comparison of Chinese versus Caucasian (and urban versus rural Chinese) were used to generate means for and percent differences in hormone concentrations. Percent difference was calculated as $(\exp^{\beta} - 1) \times 100$, where β pertains to the ethnic/racial comparison, and 95% confidence intervals (95% CI) were calculated as $\exp^{(\beta - 1 \pm 1.96[\mathrm{SE}(\beta)])} \times 100$. Means are geometric (exponentiated from the logarithmic scale). Statistical significance was defined as $P < 0.05$ (two-sided test).

Results

The mothers of Chinese neonates were substantially different from the U.S. mothers on a large number of characteristics. Chinese mothers were significantly

Table 1. Maternal, gestational, and perinatal characteristics among Caucasian women (Boston, MA) and Chinese women (Shanghai, China)

	United States		China		*P*
	n	Mean (range) or %	*n*	Mean (range) or %	
Maternal characteristics					
Age(y)	111	31.2 (19 39)	121	24.9 (20 37)	<0.0001
Weight before pregnancy (kg)	109	60.3 (43 95)	120	51.2 (38 72)	<0.0001
Height (cm)	111	164 (150 183)	121	160 (149 174)	<0.0001
BMI (kg/m^2)	109	22.4 (18 36)	120	20.0 (14 26)	<0.0001
Multiparous	48	43.2	3	2.5	<0.0001
Primiparous	63	56.8	118	97.5	
Gestational characteristics					
Gestational length (wk)	106	40.1 (37 44)	119	39.8 (34 46)	0.31
Weight gained at week 27 (kg)	106	11.5 (3.2 25)	118	9.0 (3.0 to 22)	<0.0001
Nausea and/or vomiting during pregnancy	87	78.4	77	63.6	0.02
Neither nausea or vomiting during pregnancy	24	21.6	42	34.7	
Perinatal characteristics					
Birth weight (g)	111	3,552 (2,625 4,970)	121	3,463 (1,900 4,800)	0.18
Birth length (cm)	111	50.6 (44 56)	121	49.8 (33 56)	0.10
Head circumference (cm)	109	34.8 (32 38)	119	34.7 (30 51)	0.11
Placenta weight (g)	105	586 (340 1,020)	116	633 (360 1,000)	0.005
Female	58	52.3	51	42.2	0.12
Male	53	47.7	70	57.8	

NOTE: Mean (range) for continuous variables; % for categorical variables.

younger than the U.S. mothers and, on average, were shorter, weighed less before pregnancy, had a lower body mass index (BMI), and gained less weight by the end of the second trimester (Table 1). Chinese mothers were significantly more likely to be primiparous and to have had a C-section (43.3% versus 26.1%, respectively; $P = 0.009$) compared with U.S. mothers and less likely to have completed high school (7.4% versus 97.3%; $P < 0.0001$). In addition, during the pregnancy, Chinese mothers were less likely to experience nausea and/or vomiting (63.6% versus 78.4%; $P = 0.02$), to have drunk coffee (2.5% versus 62.2%; $P < 0.0001$) or tea (12.4% versus 55.9%; $P = 0.0001$), or to have taken antibiotics (3.3% versus 22.5%; $P < 0.0001$). Alcohol consumption during pregnancy was rare among Chinese and U.S. mothers (0.8% versus 3.6%, respectively; $P_{\text{difference}} = 0.20$). There were no statistically significant differences in gender or birth size, including weight, length, and head circumference by race/ethnicity, but placental weight was significantly higher in the Chinese pregnancies. None of the 111 U.S. infants had a birth weight less than 2,500 g or a gestational length less than 37 weeks, and only 3 and 6, respectively, of the 121 Chinese infants were in these categories. There were no statistically significant differences between the Chinese and the U.S. infants, comparing the proportions in the lowest quartiles for anthropometric measurements (based on the distribution in the entire study group). The proportion of Chinese

Table 2. Distributions of cord serum hormone concentrations in U.S. Caucasians (Boston, MA) and Chinese (Shanghai, China)

Hormone	*n*	Mean	5%	25%	Median	75%	95%
Estradiol (nmol/L)							
U.S.	87	34.4	11.7	18.7	28.1	42.3	80.2
Chinese	110	44.9	3.6	15.9	36.6	57.7	125.2
Estriol (nmol/L)							
U.S.	96	541	195	354	460	664	1,135
Chinese	120	829	304	566	814	1,022	1,554
Androstenedione (nmol/L)							
U.S.	88	18.2	7.9	12.7	16.5	21.4	34.3
Chinese	114	43.4	10.9	18.4	28.0	54.9	141
Testosterone (nmol/L)							
U.S.	88	1.1	0.47	0.70	0.92	1.2	2.5
Chinese	115	5.5	0.63	1.5	3.8	7.9	17.3
Prolactin (μg/L)							
U.S.	81	334	145	249	312	436	553
Chinese	47	293	124	200	283	400	487
IGF I (nmol/L)							
U.S.	51	10.0	4.1	6.0	10.6	13.2	16.2
Chinese	22	11.8	4.4	7.0	11.3	15.4	19.2
IGFBP 3 (nmol/L)							
U.S.	52	32.0	22.3	26.9	31.6	36.3	43.7
Chinese	21	42.7	21.8	27.8	30.5	51.5	89.1

Table 3. Unadjusted means for and percent differences (95% CI) in cord hormone concentrations between U.S. Caucasians (Boston, MA) and Chinese (Shanghai, China)

Hormone	U.S., *n* (mean)	Chinese, *n* (mean)			Chinese vs U.S., % difference (95% CI)	Chinese rural vs urban, % difference (95% CI)
		Total	Urban	Rural		
Estradiol (nmol/L)	87 (28.5)	110 (29.1)	50 (28.7)	60 (29.4)	2.0 (−21.3, 32.3)	2.3 (−32.5, 55.0)
Estriol (nmol/L)	96 (477)	120 (709)	54 (663)	66 (749)	48.6 (23.6, 78.6)	13.0 (−15.3, 50.6)
Androstenedione (nmol/L)	88 (16.7)	114 (32.0)	48 (29.6)	66 (33.8)	91.9 (61.1, 128)	14.3 (−13.5, 50.9)
Testosterone (nmol/L)	88 (0.96)	115 (3.4)	51 (2.6)	64 (4.4)	257 (184, 349)	70.7 (19.4, 144)
Prolactin (μg/L)	81 (311)	47 (265)	14 (244)	33 (274)	−14.9 (−27.0, −0.92)	12.4 (−16.3, 50.8)
IGF I (nmol/L)	51 (9.1)	22 (10.5)	8 (14.2)	14 (8.8)	14.7 (−9.6, 45.4)	−37.8 (−58.3, −7.2)
IGFBP 3 (nmol/L)	52 (31.2)	21 (37.8)	8 (50.2)	13 (31.8)	21.1 (3.5, 41.8)	−36.6 (−56.5, −7.6)

NOTE: From linear regression models with logarithm-transformed hormones as the dependent variable and Chinese versus U.S. sample (or rural versus urban) as an indicator variable. Percent difference is calculated as $(\exp^{\beta} - 1) \times 100$, where β pertains to ethnic/racial comparison. Geometric means are presented.

infants in the lowest quartile of birth weight was 28.1% versus 21.6% for U.S. infants ($P = 0.29$ for difference in proportions), of head circumference was 15.1% versus 22.9% ($P = 0.18$), and of birth length was 22.3% versus 30.6% ($P = 0.17$), respectively.

The distributions of the hormones are presented in Table 2. Cord serum estriol, androstenedione, testosterone, and IGFBP-3 concentrations were significantly higher in the Chinese than in the U.S. samples (Table 3), whereas prolactin levels were significantly lower. There was no appreciable difference in estradiol concentrations by race/ethnicity, and higher IGF-I levels in the Chinese were not statistically significantly different from U.S. values. Repeating the hormone comparisons between the U.S. and the Chinese samples using Wilcoxon rank-sum tests, the results were similar, except for IGFBP-3, which did not show a statistically significant difference. The patterns of results for percent differences in cord estrogen and androgen concentrations between rural and urban Chinese were in the same direction but of lesser magnitude than those between the U.S. and all Chinese samples combined. However, only the percent difference in testosterone was statistically significant. The sample sizes, especially for prolactin, IGF-I, and IGFBP-3, were small and the 95% CIs were wide.

To determine whether differences in cord hormones between Chinese and U.S. samples were independent of the other hormones, we repeated the analyses adding each of the hormones individually to the models. With adjustment for androgen concentrations, the higher estriol and IGFBP-3 levels observed in the Chinese were attenuated. For example, with androstenedione in the model, the percent difference in estriol decreased from ~49% higher in the Chinese to only 2.2%, and adjusting IGFBP-3 for androstenedione decreased the percent difference from 21% to ~0%. In contrast, the higher androgen levels in the Chinese compared with Caucasians were not affected by adjustment for either of the estrogens or IGFBP-3. Estradiol levels did not differ between Chinese and Caucasians regardless of whether they were unadjusted or adjusted for the other hormones (data not shown).

With adjustment for gestational length, maternal age, pre-pregnancy BMI, and pregnancy weight gain through the second trimester, the androgens and IGFBP-3 remained elevated in the Chinese, but the differences in estriol and prolactin no longer remained statistically significant (Table 4). Percent differences in estriol, prolactin, and the androgens were influenced by adjustment for maternal age, and further adjustment for weight gain also affected percent differences in estriol, prolactin, and IGFBP-3 (data not shown). Estradiol concentrations, which did not differ by race/ethnicity in the unadjusted comparisons, became lower in the Chinese with adjustment (mainly from maternal age), although the racial/ethnic difference did not reach statistical significance. IGF-I levels did not differ by race/ethnicity in either unadjusted or adjusted comparisons. Additional adjustment for maternal height did not change the estimates (data not shown). Adjustment for maternal age was difficult because the Chinese women were quite young and the U.S. women were considerably older with the only appreciable overlap between 24 and 35 years old.

Table 4. Adjusted means for and percent differences (95% CI) in cord serum hormone concentrations between U.S. Caucasians (Boston, MA) and Chinese (Shanghai, China)

Hormone	*n* (Mean)	% difference*
Estradiol (nmol/L)		
U.S.	79 (34.7)	−29.8 (−52.3, 3.5)
Chinese	105 (24.4)	
Estriol (nmol/L)		
U.S.	87 (510)	27.7 (−3.1, 68.3)
Chinese	115 (651)	
Androstenedione (nmol/L)		
U.S.	80 (18.3)	60.5 (23.5, 109)
Chinese	109 (29.4)	
Testosterone (nmol/L)		
U.S.	80 (1.1)	185 (101, 304)
Chinese	110 (3.2)	
Prolactin (μg/L)		
U.S.	73 (301)	−10.1 (−30.6, 16.3)
Chinese	47 (270)	
IGF I (nmol/L)		
U.S.	45 (9.0)	11.7 (−23.6, 63.2)
Chinese	22 (10.0)	
IGFBP 3 (nmol/L)		
U.S.	46 (29.4)	40.4 (7.8, 82.9)
Chinese	21 (41.3)	

NOTE: From linear regression models with logarithm-transformed hormones as the dependent variable and Chinese versus U.S. samples as an indicator variable. Percent difference is calculated as $(\exp^{\beta} - 1) \times 100$, where β pertains to ethnic/racial comparison. Geometric means are presented.

*Model includes gestational length, maternal age, pre-pregnancy BMI, and weight gain as independent variables.

Repeating the analyses in this age range (24-35 years) and controlling for age as a continuous variable, the results were similar to those observed in the overall group (data not shown). Additional adjustment for placental weight did not change the results, although the higher estriol concentrations among Chinese neonates increased from ~28% to 38% (95% CI, 2.5%, 85%). Furthermore, adding offspring gender to the adjusted models did not change the percent differences in the overall analysis (estradiol, −31.8% with offspring gender versus −29.8% without offspring gender; estriol, 24.4% versus 27.7%; androstenedione, 53.8% versus 60.5%; testosterone, 166% versus 185%; prolactin, −10.7% versus −10.1%; IGF-I, 14.3% versus 11.7%; IGFBP-3, 41.7% versus 40.4%).

Nearly all of the Chinese women were primiparous; therefore, we restricted analyses to neonates born to women with no previous live or still births. The pattern of unadjusted and adjusted results was generally similar to the overall differences with significantly elevated estriol, androstenedione, testosterone, and IGFBP-3 concentrations in the Chinese. For example, the unadjusted and adjusted (for maternal age, pre-pregnancy weight, and weight gain) percent differences in estriol in all women were 48.6% and 27.7% versus 47.5% and 22.0% in primiparous women. The corresponding values were 91.9% and 60.5% versus 78.5% and 56.4% for androstenedione and 257% and 185% versus 225% and 162% for testosterone. For estradiol, the unadjusted values were similar, whereas the adjusted differences became somewhat greater and achieved statistical significance [unadjusted percent difference, −10.4% (95% CI, −35.4%, 24.2%); adjusted percent difference, −45.2% (95% CI, −65.2%, −13.6%)]. As the results for hormones by race/ethnicity were in the same direction and remained statistically significant regardless of parity, we included multiparous women in the overall analysis to increase statistical power. When stratified by offspring gender, the magnitude of the percent differences was generally greater in males than females, but these comparisons were limited by the reduction in sample sizes. For example, the percent differences were 98.3% (95% CI, 53.7%, 156%) for androstenedione and 276% (95% CI, 178%, 408%) for testosterone in the male infants and 74.5% (95% CI, 38.8%, 120%) and 213% (95% CI, 123%, 339%) among the females, respectively. None of the interactions of the associations of hormones and race/ethnicity by offspring gender were statistically significant, ranging from 0.20 for IGF-I to 0.91 for prolactin.

Discussion

The data presented here are the first to directly address differences in fetal hormone concentrations between pregnancies occurring in geographic regions characterized by relatively low (China) and high (United States) breast cancer incidence rates. The elevated cord androgen concentrations we observed, however, are consistent with previous studies of pregnancies occurring in North America showing higher dehydroepiandrosterone sulfate (DHEAS) concentrations in Chinese Canadians compared with Caucasians (20), as is the pattern of higher cord estriol concentrations that we observed in the Chinese (20). In contrast, estradiol concentrations were not elevated in the Chinese neonates in our study. In a previous study, cord estradiol concentrations were higher in Chinese American than in Caucasian pregnancies (21) with or without adjustment for several pregnancy factors, although the study was small and consisted of women recruited from clients of a cord blood registry, suggesting relatively high acculturation. We found no difference in IGF-I by race/ethnicity in our data, whereas IGFBP-3 levels were elevated in the Chinese. The study cited above (21) reported no differences in unadjusted IGF profiles, including IGF-I and IGFBP-3 levels, comparing Chinese American with U.S. pregnancies.

Our findings for differences in cord estrogen and prolactin levels between Chinese and U.S. neonates were not entirely similar to hormone differences observed in the maternal data for these same pregnancies. Whereas androgen and IGF concentrations were not assessed in the study of serum hormones in the mothers of these infants (13), maternal levels of estradiol and prolactin as well as estriol were elevated in the Chinese. The vast majority (over 90%) of estradiol and estriol enter the maternal compartment from the placenta (22). Most studies have measured hormones and other biomarkers in the maternal circulation due to the difficulty in directly sampling the *in utero* environment and based on the assumption that maternal hormones and other endocrine factors reflect those in the fetal circulation because of the highly integrated maternal-placental-fetal unit. However, the degree of correlation between, for example, estrogen and androgen concentrations in the maternal and fetal circulations is modest (23) and may explain the differences in the direction of results for the maternal and cord samples in the present data.

We are not aware of studies that have measured fetal androgen concentrations at multiple points throughout the pregnancy. Whether hormone levels in the fetal circulation reflect levels in breast tissue is not known and is an issue in any study that uses such proxies for fetal exposure. In epidemiologic studies, umbilical cord sampling is only feasible after delivery; thus, the differences we observed between groups may not represent real differences earlier in pregnancy. In particular, androgen concentrations may be higher after vaginal deliveries than after C-sections. However, the proportion of C-sections was actually greater in the Chinese than U.S. women (41.3% versus 26.1%) and thus would not explain the higher androgen concentrations we observed in the Chinese cord samples.

We observed elevated androgen and estriol concentrations but either no difference (unadjusted) or reductions (adjusted) in estradiol in the Chinese. In uncomplicated pregnancies, nearly comparable amounts of DHEAS from the maternal and fetal adrenal glands are enzymatically converted in the placenta to androstenedione and testosterone, which are then aromatized to estrone and estradiol, respectively (22). The conversion rate of androgens to estradiol is a function of placental size and capacity as well as of aromatase enzyme levels. Estriol is synthesized by the placenta from 16α-hydroxy-DHEAS, which is formed in the fetal liver from DHEAS. Over 90% of urinary estriol are ultimately derived from the fetal adrenal gland (22). As pregnancy estriol is derived from androgen substrate, these hormones are necessarily correlated (r = 0.39 for androstenedione and estriol and r = 0.30 for testosterone and estriol in the present study). Thus, the higher estriol concentrations in the Chinese that

we observed could merely be due to higher androgen concentrations, which were also noted in the Chinese. When androstenedione was added to the regression model for estriol, the higher estriol concentrations in the Chinese infants were markedly attenuated, whereas androstenedione remained significantly higher in the Chinese. This suggests that the elevated estriol levels in the Chinese are probably due to their greater amounts of estrogen precursor (that is, fetal androgens). The attenuation in the group differences for estriol with androstenedione in the model could also result if androstenedione was measured with less laboratory error than estriol. However, the coefficients of variation were fairly similar for the two hormones.

The similar or reduced estradiol levels in the presence of elevated testosterone concentrations are more difficult to explain. This implies less aromatization in the Chinese, but it does not appear to be explained by placental size as average placental weight was higher in the Chinese pregnancies. Incomplete aromatization in the placental compartment could also explain the higher testosterone concentrations in the Chinese. Alternatively, or in addition to this explanation, the higher androgens in the Chinese may be due to greater concentrations in the fetal compartment. In this regard, the hormone results by offspring gender may add to the understanding of the biology of the higher androgen levels in the Chinese. Given the values are higher in the Chinese regardless of gender implies that placental differences (that is, in aromatization) between Chinese and U.S. pregnancies may be responsible.

Our unadjusted data would not appear to support the hypothesis that exposure to lower estrogen levels *in utero* are responsible in part for the lower breast cancer risk in Chinese women. Other prenatal factors related to breast cancer risk have been proposed to be mediated through differences in pregnancy estrogen exposure, including high birth weight (24) and dizygotic twinning (25-30) as well as the reduced risk observed for women born of preeclamptic pregnancies (31). Birth weight has been positively associated with maternal estrogens in several studies (32-34). However, data validating the associations of preeclampsia (34, 35) and birth weight (21, 36-38) with estrogen levels, particularly in the cord, are conflicting, and data for dizygotic twinning are lacking (39).

We hypothesize that a difference in fetal androgen exposure at a critical period during pregnancy may explain the lower breast cancer rates in Asians compared with Caucasians. Androgen concentrations in the cord circulation of the Chinese neonates were two to three times greater than in the U.S. neonates. Elevated fetal androgen concentrations have been proposed as mediating the associations of prenatal exposures with breast cancer risk (40) possibly through reduction of the initial breast stem cell population. Suppression of embryonic mammary gland development by androgens in males supports this hypothesis (41). In the mouse model, destruction of mammary gland anlagen (the initial clustering of cells destined to become breast tissue) by testosterone occurs in early pregnancy, and this androgen sensitivity is expressed in male as well as female mammary glands (42). Female fetuses are protected from sterilization through rapid placental metabolism of androgens to estrogens (43). Given that the prohibitive effects of androgens on breast anlagen formation occur in female as well as male mammary glands and the empirical evidence that women undergo breast development whereas men generally do not, we believe rapid androgen metabolism plays a protective role in allowing female breast development as well as in preventing virilization. Androgen variability below the level of sterilization, however, may have significant biological consequences. Thus, it is possible that limited androgen transfer back to the fetus in cases in which androgen levels are high could influence anlagen formation, as it does in males. Whether the magnitude of the difference in androgen concentrations between Chinese and U.S. infants that we observed is sufficient to protect the breast is unknown. Elevated androgen concentrations are observed for other prenatal exposures that are associated with a reduced breast cancer risk, such as preeclampsia (34, 35). These observations would be consistent with a protective effect of fetal androgens but the data, particularly for cord concentrations, are sparse.

Associations of the hormones studied with other maternal and pregnancy factors could explain the differences observed by race/ethnicity. In particular, the Chinese mothers tended to be younger and physically smaller with lower pre-pregnancy weight, height, and BMI and less pregnancy weight gain. The differences in androgens and IGFBP-3 remained with adjustment for these factors, and those for estriol and prolactin were attenuated. In contrast, the estradiol concentrations became lower than in Caucasians following adjustment. The similar estradiol concentrations in the two racial/ethnic groups in unadjusted comparisons were largely due to younger maternal age in Chinese women.

We presented both unadjusted means and those adjusted for maternal and pregnancy factors shown previously to be associated with cord hormone levels and which were related to race/ethnicity in these data. We believe the unadjusted results address whether cord hormones explain the international difference in breast cancer rates in offspring, rate differences that are unadjusted for maternal characteristics. For example, if estradiol concentrations are similar in the Chinese and U.S. mothers because the Chinese mothers tend to be younger and smaller, then (assuming our population samples are representative and Chinese women are indeed generally younger and smaller) estradiol seems unlikely to explain the international rate differences. The unadjusted results also could be more relevant to the actual pregnancy exposure of the fetus. In subsequent analyses, we adjusted for maternal factors, including age, pre-pregnancy BMI and weight gain, and length of the gestation.

The adjusted results more appropriately address whether the hormone differences we observed between Chinese and U.S. infants are due to differences in maternal age and size. Some but not all of the higher androgen levels in the Chinese appear to be due to their younger maternal age. For estriol and prolactin, the higher levels in the Chinese appear to be due to their younger maternal age and to less pregnancy weight gain. The lack of difference in the crude data for estradiol seems to be driven by the younger maternal age of the Chinese. If either androgen or estrogen levels are causally responsible for the international differences in breast cancer risk, it would seem that their correlates (that is, maternal age) would be risk factors for breast

cancer. However, until now, this has not been consistently observed (44). Regardless of which results are used, the unadjusted or adjusted, and focusing only on fetal exposure, higher androgens and IGFBP-3 are consistent with protective effects on breast cancer risk, whereas prolactin is consistent with adverse effects.

The samples of women from China and the United States were not population based, and as such, the data may not be representative of pregnancies occurring in each country. The Chinese women, however, were mainly from Shanghai and likely had a more western lifestyle than would be typical in China. Thus, any observed hormone differences that were due to environmental factors may be underestimated. In fact, the hormone profile for the urban Chinese, although closer to the rural Chinese, was between that of the Caucasians and that of rural Chinese. Some of the cord sera showed signs of hemolysis and could not be used for the direct assays (that is, prolactin, IGF-I, and IGFBP-3), resulting in smaller sample sizes for these analytes and less power to detect differences by race/ethnicity.

In conclusion, we found higher concentrations of estriol, androstenedione, testosterone, and IGFBP-3 and lower prolactin in cord serum from Chinese compared with U.S. neonates, whereas levels of estradiol and IGF-I were not different. These data are consistent with the hypothesis that elevated prenatal androgen exposure may be protective against subsequent breast carcinogenesis. Whereas the focus of this paper is on breast cancer, these results may be relevant for other endocrine cancers that have been suggested as being associated with fetal hormone exposure, including testicular and prostate cancer. Given that cultural and lifestyle practices are changing in parts of Asia, studying populations, for example, that have moved from rural to urban areas to address whether changes in lifestyle factors affect pregnancy-maternal and umbilical cord-hormone concentrations could be useful.

References

1. CANCER*Mondial*; 2002. Available from: http://www-dep.iarc.fr/.
2. Ziegler RG, Hoover RN, Pike MC, et al. Migration patterns and breast cancer risk in Asian-American Women. J Natl Cancer Inst 1993;85:1819 27.
3. Wu AH, Ziegler RG, Pike MC, et al. Menstrual and reproductive factors and risk of breast cancer in Asian-Americans. Br J Cancer 1996;73:680 6.
4. Ursin G, Wu AH, Hoover RN, et al. Breast cancer and oral contraceptive use in Asian-American women. Am J Epidemiol 1999;150:561 7.
5. Buell P. Changing incidence of breast cancer in Japanese-American women. J Natl Cancer Inst 1973;51:1479 83.
6. Miller AB, Bulbrook RD. UICC Multidisciplinary Project on Breast Cancer: the epidemiology, aetiology and prevention of breast cancer. Int J Cancer 1986;37:173 7.
7. Willett WC, MacMahon B. Diet and cancer an overview. N Engl J Med 1984;310:633 8, 697 703.
8. National Cancer Institute. Summary report: early reproductive events and breast cancer workshop; 2003. Available from: http://cancer.gov/cancerinfo/ereworkshop-report.
9. Trichopoulos D. Hypothesis: does breast cancer originate *in utero*? Lancet 1990;335:939 40.
10. Key TJA, Chen J, Wang DY, Pike MC, Boreham J. Sex hormones in women in rural China and in Britain. Br J Cancer 1990;62:631 6.
11. Bernstein L, Yuan J-M, Ross RK, et al. Serum hormone levels in premenopausal Chinese women in Shanghai and white women in Los Angeles: results from two breast cancer case-control studies. Cancer Causes Control 1990;1:51 8.
12. Falk RT, Fears TR, Hoover RN, et al. Does place of birth influence endogenous hormone levels in Asian-American women. Br J Cancer 2002;87:54 60.
13. Lipworth L, Hsieh C-C, Wide L, et al. Maternal pregnancy hormone levels in an area with a high incidence (Boston, USA) and in an area with a low incidence (Shanghai, China) of breast cancer. Br J Cancer 1999;79:7 12.
14. Lambe M, Trichopoulos D, Hsieh CC, Wuu J, Adami HO, Wide L. Ethnic differences in breast cancer risk: a possible role for pregnancy levels of alpha-fetoprotein? Epidemiology 2003;14:85 9.
15. Goebelsmann U, Bernstein GS, Gale JA, et al. Serum gonadotropin, testosterone, estradiol and estrone levels prior to and following bilateral vasectomy. In: Lepow IH, Crozier R, eds. Vasectomy: immunologic and pathophysiologic effects in animals and man. New York: Academic Press; 1979. p. 165.
16. Goebelsmann U, Arce JJ, Thorneycroft IH, Mishell DR, Jr. Serum testosterone concentrations in women throughout the menstrual cycle and following hCG administration. Am J Obstet Gynecol 1974;119:445 52.
17. Goebelsmann U, Horton R, Mestman JH, et al. Male pseudohermaphroditism due to testicular 17-hydroxysteroid dehydrogenase deficiency. J Clin Endocrinol Metab 1973;36:867.
18. Katagiri H, Stanczyk FZ, Goebelsmann U. Estriol in pregnancy. III. Development, comparison and use of specific antisera for rapid radioimmunoassay of unconjugated estriol in pregnancy plasma. Steroids 1974;24:225 38.
19. Armitage P, Berry G. Statistical methods in medical research. Boston: Blackwell Boston; 1987.
20. Yeung LPK, Wong ACK, Wang X, Birmingham CL, Lewicka S, Chanoine J-P. Different relationship between anthropometric markers and umbilical cord plasma leptin in Asian and Caucasian neonates. Pediatric Res 2003;53:1019 24.
21. Shibata A, Harris DT, Billings PR. Concentrations of estrogens and IGFs in umbilical cord blood plasma: a comparison among Caucasian, Hispanic, and Asian-American females. J Clin Endocrinol Metab 2002;87:810 5.
22. Carr BR. The maternal-fetal-placental unit. In: Becker KL, editor. Principles and practice of endocrinology and metabolism. 2nd ed. Philadelphia: J.B. Lippincott; 1995.
23. Troisi R, Potischman N, Roberts JM, et al. Correlation of serum hormone concentrations in maternal and umbilical cord samples. Cancer Epidemiol Biomarkers Prev 2003;12:452 6.
24. Lagiou P, Samoli E, Lagiou A, Hsieh CC, Adami HO, Trichopoulos D. Maternal height, pregnancy estriol and birth weight in reference to breast cancer risk in Boston and Shanghai. Int J Cancer 2005;117:494 8.
25. Ekbom A, Hsieh C-C, Lipworth L, Adami H-O, Trichopoulos D. Intrauterine environment and breast cancer risk in women: a population-based study. J Natl Cancer Inst 1997;88:71 6.
26. Hsieh C-C, Lan S-J, Ekbom A, Petridou E, Adami H-O, Trichopoulos D. Twin membership and breast cancer risk. Am J Epidemiol 1992;136:1321 6.
27. Weiss HA, Potischman NA, Brinton LA, et al. Prenatal and perinatal risk factors for breast cancer in young women. Epidemiology 1997;8:181 7.
28. Braun MM, Ahlbom A, Floderus B, Brinton L, Hoover RN. Effect of twinship on incidence of cancer of the testis, breast and other sites (Sweden). Cancer Causes Control 1995;6:519 24.
29. Cerhan JR, Kushi LH, Olson JE, et al. Twinship and risk of postmenopausal breast cancer. J Natl Cancer Inst 2000;92:261 5.
30. Verkasalo PK, Kaprio J, Pukkala E, Koskenvuo M. Breast cancer risk in monozygotic and dizygotic female twins: a 20-year population-based cohort study in Finland from 1976 to 1995. Cancer Epidemiol Biomarkers Prev 1999;8:271 4.
31. Innes KE, Byers TE. Preeclampsia and breast cancer risk. Epidemiology 1999;10:722 32.
32. Mucci LA, Lagiou P, Tamimi RM, Hsieh CC, Adami H-O, Trichopoulos D. Pregnancy estriol, estradiol, progesterone and prolactin in relation to birth weight and other birth size variables (United States). Cancer Causes Control 2003;14:311 8.
33. Kaijser M, Granath F, Jacobsen G, Cnattingius S, Ekbom A. Maternal pregnancy estriol levels in relation to anamnestic and fetal anthropometric data. Epidemiology 2000;11:315 9.
34. Troisi R, Potischman N, Neuser Johnson C, et al. Estrogen and androgen concentrations are not lower in the umbilical cord serum of preeclamptic pregnancies. Cancer Epidemiol Biomarkers Prev 2003;12:1268 70.
35. Atamer Y, Erden AC, Demir B, Koçyigit Y, Atamer A. The relationship between plasma levels of leptin and androgen in healthy and preeclamptic pregnant women. Acta Obstet Gynecol Scand 2004;83:425 30.

36. Troisi R, Potischman N, Roberts JM, Siiteri P, Hoover RN. Associations of maternal and umbilical cord hormone concentrations with maternal, gestational and neonatal factors. Cancer Causes Control 2003;14:347 55.

37. Nagata C, Iwasa S, Shiraki M, Shimizu H. Estrogen and α-fetoprotein levels in maternal and umbilical cord blood samples in relation to birth weight. Cancer Epidemiol Biomarkers Prev 2006; 15:1469 72.

38. Simmons D, France JT, Keelan JA, Song L, Knox BS. Sex differences in umbilical cord serum levels of inhibin, testosterone, oestradiol, dehydroepiandrosterone sulphate, and sex hormone-binding globulin in human term neonates. Biol Neonate 1994;65:287 94.

39. Thomas HV, Murphy MF, Key TJ, Fentiman IS, Allen DS, Kinlen LJ. Pregnancy and menstrual hormone levels in mothers of twins compared to mothers of singletons. Ann Hum Biol 1998;25:69 75.

40. Troisi R, Potischman N, Roberts JM, et al. Maternal serum hormone concentrations in preeclamptic and uncomplicated pregnancies. Int J Epidemiol 2003;32:455 60.

41. Drews U, Drews U. Regression of mouse mammary gland anlagen in recombinants of *Tfm* and wild-type tissues: testosterone acts via the mesenchyme. Cell 1977;10:401 4.

42. Wald N, Cuckle H, Wu TS, George L. Maternal serum unconjugated oestriol and human chorionic gonadotrophin levels in twin pregnancies: implications for screening for Down's syndrome. Br J Obstet Gynaecol 1991;98:905 8.

43. Yen SSC. Endocrine-metabolic adaptations in pregnancy. In: Yen SSC, Jaffe RB, editors. Reproductive endocrinology. Philadelphia: W.B. Saunders; 1991. p. 936 81.

44. Potischman N, Troisi R. *In utero* and early life exposures in relation to risk of breast cancer. Cancer Causes Control 1999;10:561 73.

Breast Cancer RESEARCH

RESEARCH ARTICLE Open Access

Genetic variation in the estrogen metabolic pathway and mammographic density as an intermediate phenotype of breast cancer

Jingmei Li[1,2*], Louise Eriksson[1], Keith Humphreys[1], Kamila Czene[1], Jianjun Liu[2], Rulla M Tamimi[3,4], Sara Lindström[5], David J Hunter[3], Celine M Vachon[6], Fergus J Couch[7], Christopher G Scott[6], Pagona Lagiou[4,8], Per Hall[1]

Abstract

Introduction: Several studies have examined the effect of genetic variants in genes involved in the estrogen metabolic pathway on mammographic density, but the number of loci studied and the sample sizes evaluated have been small and pathways have not been evaluated comprehensively. In this study, we evaluate the association between mammographic density and genetic variants of the estrogen metabolic pathway.

Methods: A total of 239 SNPs in 34 estrogen metabolic genes were studied in 1,731 Swedish women who participated in a breast cancer case control study, of which 891 were cases and 840 were controls. Film mammograms of the medio lateral oblique view were digitalized and the software Cumulus was used for computer assisted semi automated thresholding of mammographic density. Generalized linear models controlling for possible confounders were used to evaluate the effects of SNPs on mammographic density. Results found to be nominally significant were examined in two independent populations. The admixture maximum likelihood based global test was performed to evaluate the cumulative effect from multiple SNPs within the whole metabolic pathway and three subpathways for androgen synthesis, androgen to estrogen conversion and estrogen removal.

Results: Genetic variants of genes involved in estrogen metabolism exhibited no appreciable effect on mammographic density. None of the nominally significant findings were validated. In addition, global analyses on the overall estrogen metabolic pathway and its subpathways did not yield statistically significant results.

Conclusions: Overall, there is no conclusive evidence that genetic variants in genes involved in the estrogen metabolic pathway are associated with mammographic density in postmenopausal women.

Introduction

Mammographic breast density is one of the strongest risk factors for breast cancer. Several studies have shown that women with extensive dense tissue are at two to six times higher risk of developing the disease than women of similar age with lower mammographic density [1,2]. A strong genetic basis has been suggested for mammographic density [3]. Twin studies have estimated the heritability of this trait to be between 60 and 67% [4]. Evidence for a genetic influence also comes from other studies on family history, familial aggregation and segregation analyses [5,6].

Mammographic density is strongly correlated with hormone exposure profiles of women [7]. Several hormonal risk factors for breast cancer have been found to influence mammographic density in a similar fashion to their respective associations with risk for the disease [8]. For example, a strong inverse relationship has been observed between parity on mammographic density [9]. In addition, hormone replacement therapy (HRT) users and women who have a late first-born child or late menopause have higher breast densities on average [9]. In view of evidence suggesting an association between mammographic density and hormone-related factors, and the fact that estrogen is a strong risk factor for postmenopausal breast cancer, efforts have been made to identify underlying genetic determinants of

* Correspondence: Jingmei.Li@ki.se
[1]Karolinska Institutet, Department of Medical Epidemiology and Biostatistics, Box 281, 171 77 Stockholm, Sweden

mammographic density within pathways related to steroid hormone biosynthesis and metabolism [10-13]. Such endeavors assume mammographic density to be an intermediate phenotype for breast cancer. Several genes involved in hormone-related pathways - such as HSD3B1 [5,14], COMT [11,14] and ESR1 [15] - have been suggested to be associated with mammographic breast density. Findings are inconsistent, however, and only few candidate genes have been studied at a time.

We recently reported the results of a study evaluating a total of 239 SNPs in 34 estrogen metabolic genes in 1,596 breast cancer cases and 1,730 population controls from Sweden, of which the outcome variable was breast cancer (Low *et al*., manuscript submitted). No significant SNP association was evident after correction for multiple testing, but pathway-based global tests revealed significant association evidence for the overall estrogen metabolic pathway (P = 0.034) and, in particular, the androgen-to-estrogen conversion subpathway (P = 0.007). In the present study, we comprehensively examine genetic variation in the estrogen metabolic pathway with mammographic density. The number of SNPs and genes studied provides the most extensive coverage to date with respect to studying mammographic breast density.

Materials and methods

Study subjects

The subjects included in the current study are drawn from a population-based case-control study of postmenopausal breast cancer in women born in Sweden aged 50 to 74 years at the time of enrollment, which was between 1 October 1993 and 31 March 1995. Controls were randomly selected from the Swedish Total Population Register and were frequency matched to the expected age distribution of the cases. Details on data collection and subjects have been described previously [16]. The final study group with both mammographic density and genotype data included 891 breast cancer cases and 840 controls. Although all women were postmenopausal at the time of recruitment to the parent study, a subset of the women (43/1,731) was premenopausal in reference to the date of mammogram.

Approval of the study was given by the ethical review board at the Karolinska Institutet (Stockholm, Sweden) and six other ethical review boards in the respective regions in which the subjects were based, and informed consent was obtained from each participant.

Validation of SNPs with significant associations was performed using mammographic density data from two other studies.

Mammographic density data

The process of collecting mammographic density data in this study has been described previously [17]. Film mammograms of the medio-lateral oblique view were digitized using an Array 2905HD Laser Film Digitizer (Array Corporation, Tokyo, Japan), which covers a range of 0 to 4.7 optical density. For controls, the breast side was randomized. For cases, the side contralateral to the tumor was used. The density resolution was set at 12-bit spatial resolution. The Cumulus software used for the computer-assisted measurement was developed at the University of Toronto [18]. For each image, a trained observer (LE) set the appropriate gray-scale threshold levels defining the edge of the breast and distinguishing dense from nondense tissue. The software calculated the total number of pixels within the entire region of interest and within the region identified as dense. These values were used to calculate the percentage of the breast area that is dense. A random 10% of the images were included as replicates to assess the intra-observer reliability, which was high with a Spearman rank correlation coefficient of 0.95.

Gene and SNP selection

The process of gene and SNP selection has been described in detail by Low *et al*. (manuscript submitted). A total of 1,007 SNPs were selected from 35 genes and their 30 kb flanking sequences that code the enzymes involved in estradiol or estrone metabolism and are expressed in the breast. These SNPs were genotyped in 92 Swedish control samples to assess linkage disequilibrium patterns, to select tagging SNPs (tagSNPs) and to evaluate their coverage.

Haplotypes were reconstructed using the partition-ligation-expectation-maximization algorithm [19] implemented in the *tagSNPs* program [20]. A subset of tagSNPs were selected based on the R^2 coefficient, which quantifies how well the tagSNP haplotypes predict the genotype or the number of copies of haplotypes an individual carries. The performance of tagSNPs in capturing unobserved SNPs within the genes was evaluated using a SNP-dropping analysis. In brief, each of the genotyped SNPs was dropped in turn and then tagSNPs were selected from the remaining SNPs so that their haplotypes predicted the remaining SNPs with an R^2 value of 0.85. In total, 312 tagSNPs from the 35 genes were selected for genotyping.

Figure 1 delineates the processes and genes involved in the androgen synthesis, androgen-to-estrogen conversion and estrogen removal subpathways. The lists of SNPs corresponding to each subpathway are summarized in Tables S1 to S3 in Additional file 1.

DNA extraction and genotyping

DNA was extracted from 4 ml whole blood using the QIAamp DNA Blood Maxi Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions

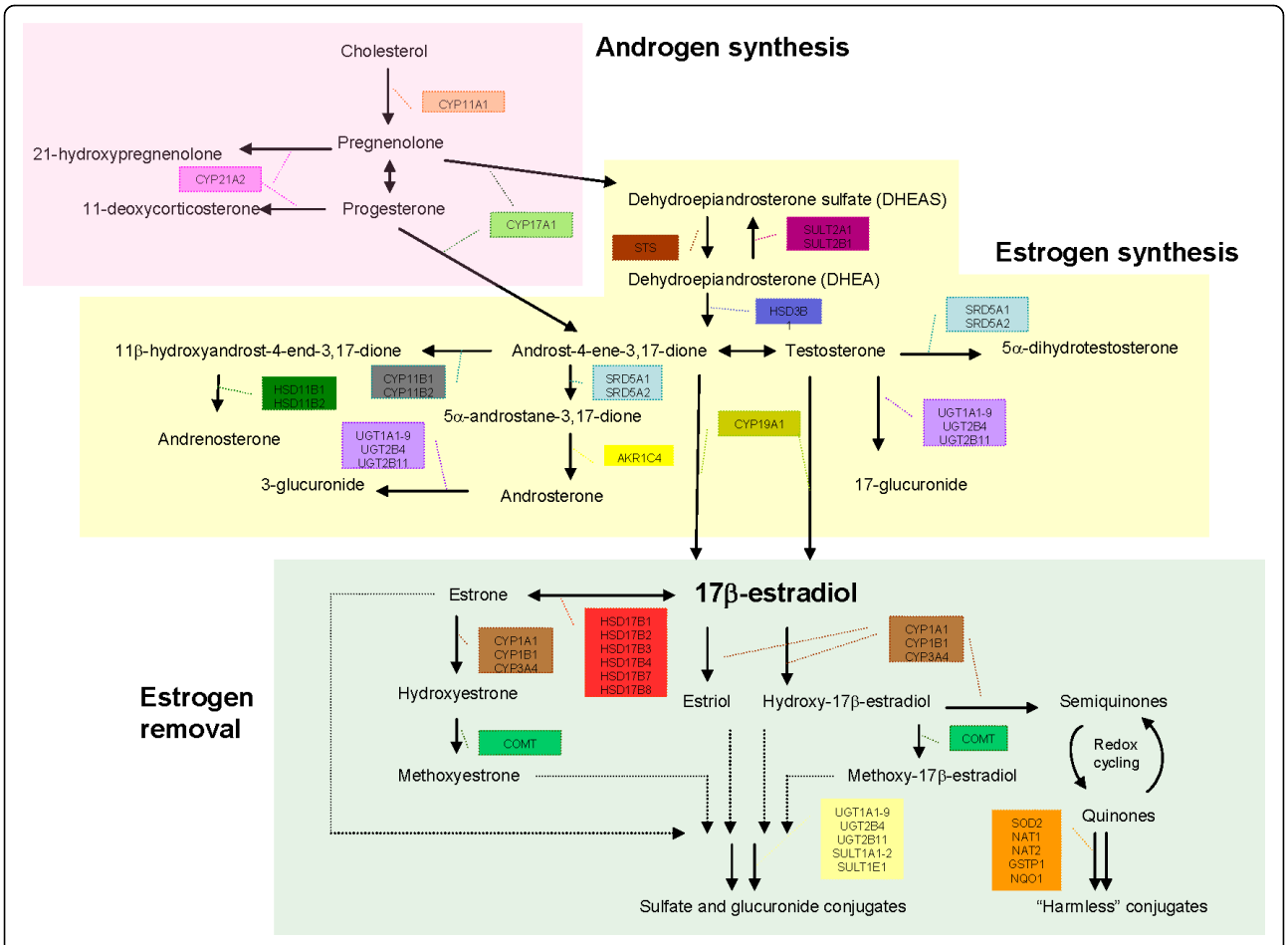


Figure 1 Subdivision of the estrogen metabolic pathway. The 34 metabolic genes analyzed in the present study are involved in different steps of the estrogen metabolism. The genes are divided into the three groups involved in androgen synthesis, estrogen synthesis and estrogen removal for subpathway based association analysis.

and nonmalignant cells in paraffin-embedded tissue using a standard phenol/chloroform/isoamyl alcohol protocol. Genotyping was performed using the primer extension-based assay from Sequenom (San Diego, CA, USA) according to the manufacturers' instructions. DNA samples were randomly assigned to the plates carrying positive and negative controls, and all genotyping results were generated and checked by laboratory staff unaware of the case-control status. SNPs with a call rate <85%, minor allele frequency <1% or out of Hardy-Weinberg equilibrium ($P < 0.05/312$) were excluded from further analysis. The genotype concordance was >99%, suggesting high genotyping accuracy. Overall, 239 tagSNPs from the 34 genes were successfully genotyped and used in statistical analysis.

Statistical analysis

Linear regression models were fitted, treating percentage density as an outcome. Models were adjusted for age, body mass index, menopausal status and HRT. Age was coded as 0, 1 and 2 for women <50 years, between 50 and 60 years, and >60 years of age, respectively. The body mass index was treated as a continuous variable. Menopausal status was determined from the time difference between the date of menopause and the date on which the mammogram was taken. HRT was considered a categorical variable made up of three groups: never users, past users and current users. The mammographic density measurements were transformed by the power of 0.3, yielding an approximately normal distribution. The genotypes were coded 0, 1 and 2 and treated as continuous variables.

A likelihood ratio test was performed for each SNP. Normal quantile-quantile plots were used to examine the distributions of the $-\log_{10}$-transformed P values. To assess whether the SNPs associated with breast cancer risk are the same SNPs as those associated with mammographic density, we used the Spearman's rank correlation test, evaluating the relationship between odds ratios corresponding to SNP effects on breast cancer

risk and the regression coefficients of SNP effects on percentage density. The admixture maximum likelihood-based global test [21] was performed to evaluate the cumulative effect on mammographic density from multiple SNPs within the whole metabolic pathway and three subpathways for androgen synthesis, androgen-to-estrogen conversion and estrogen removal. Affection status for the admixture maximum likelihood analysis was defined by taking the lowest quantile of all percentage density measurements as controls and the highest quantile as cases. *P* values of the admixture maximum likelihood test were obtained via 5,000 permutations. Software R (v2.8.0) [22] and admixture maximum likelihood [21] were used for data management, quality control and statistical analyses.

Validation of significantly associated SNPs

SNP associations with mammographic density were validated in 1,590 women genotyped with the Illumina HumanHap500 as part of the Cancer Genetic Markers of Susceptibility Project (CGEMS) [23]. The CGEMS project is a National Cancer Institute initiative to conduct genome-wide association studies to identify genes involved in breast cancer and prostate cancer. The initial CGEMS breast cancer scan was designed and funded to study the main effect of SNP variants on breast cancer risk in postmenopausal women, and has been completed [24]. Briefly, the first stage of the project involved a whole genome scan of 1,145 invasive postmenopausal breast cancer cases and 1,142 matched controls from the Nurses' Health Study nested case-control study [24]. The Nurses' Health Study was initiated in 1976, when 121,700 US registered nurses aged 30 to 55 returned an initial questionnaire [25]. During 1989 and 1990, blood samples were collected from 32,826 women [26]. For 1,590 of these women - of which 806 were breast cancer cases and 784 were healthy controls - we also had mammographic density measurements.

We collected mammograms as close as possible to the date of blood collection (1989 to 1990). To assess mammographic density, the craniocaudal (CC) views of both breasts were digitized at 261 μm/pixel with a Lumysis 85 laser film scanner, which covers a range of 0 to 4.0 optical density. The software for computer-assisted thresholding was developed at the University of Toronto [18]. We used the average percentage density of both breasts for this analysis. This collection has been described in detail in a previous publication [27]. SNPs not available on the Illumina HumanHap550 panel were imputed using MACH [28] based on HapMap Phase II (release 21a). For the analysis of imputed data, the ProbABEL package from the ABEL set of programs was used [29]. Percentage density was transformed by the power of 0.3 to be consistent with the parent study. This study was approved by the Committee on the Use of Human Subjects in Research at Brigham and Women's Hospital.

The second validation population consisted of a set of controls from an ongoing breast cancer case-control study at the Mayo Clinic. Briefly, the Mayo Clinic Breast Cancer Study is an Institutional Review Board-approved, clinic-based, case-control study initiated in February 2001 at Mayo Clinic, Rochester, MN, USA. The study design has been presented previously [30,31]. Clinic attendance formed the sampling frame for Mayo Clinic cases and controls. Consecutive cases were women aged 18 years or over with histologically confirmed primary invasive breast carcinoma and recruited within 6 months of the date of diagnosis. Cases lived in the six-state region that defines Mayo Clinic's primary service population (Minnesota, Iowa, Wisconsin, Illinois, North Dakota, and South Dakota). Controls without prior history of cancer (other than nonmelanoma skin cancer) were frequency matched on age (5-year age category), race and six-state region of residence to cases. Controls were recruited from the outpatient practice of the Divisions of General Internal Medicine and Primary Care Internal Medicine at Mayo Clinic, where they were seen for routine medical examinations.

The present analysis genotyped Caucasian controls (99% of study participants) enrolled through September 2007, who had mammograms available, representing 995 total controls (76% of total possible controls), of which 783 were postmenopausal. Screening mammograms were ascertained close to the enrollment date and the left CC view was digitized on an Array 2905HD Laser Film Digitizer, which covers a range of 0 to 4.7 optical density. Percentage mammographic density was estimated by an expert reader [32] on the left CC view, using the same Cumulus software described above [33]. Genotyping was carried out using TaqMan (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions, using 10 to 20 ng DNA. Primers and probes were Assay-by Design (Applied Biosystems). Following PCR amplification, end reactions are read on the ABI Prism 7900 ht using Sequence Detection Software (Applied Biosystems). SNP associations were examined only in the 783 postmenopausal controls, to be comparable with the two other populations. The percentage density was transformed by the power of 0.3 to be consistent with the parent study.

Results

Our dataset consisted of 1,731 postmenopausal women, of which 981 were breast cancer cases and 840 were controls (Table 1). Cases and controls differed significantly in age at first birth ($P = 0.0126$), parity ($P < 0.0001$), family history of breast cancer ($P = 0.0002$) and

Table 1 Selected characteristics of subjects

	Breast cancer cases (n = 891)		Breast cancer controls (n = 840)		
	Mean	SD	Mean	SD	*P* value
Age (years)	63.0	6.3	63.0	6.3	0.9045
Height (cm)	164.1	5.7	163.6	5.5	0.0766
Weight (kg)	68.9	110	68.8	11.6	0.8153
Body mass index	25.6	3.9	25.6	4.1	0.8420
Age at menarche (years)	13.6	1.4	13.6	1.5	0.6090
Age at first birth (years)	25.4	50	24.8	4.7	0.0126
Parity	1.9	1.2	2.2	1.3	0.0000
Age at menopause (years)	50.3	3.6	50.1	3.9	0.1223
HRT (% ever use)	0.53		0.50		0.2523
Family history (%)	0.15		0.09		0.0002
Percent density	16.7	14.3	14.6	14.0	0.0017

Means and standard deviations (SD) are given for continuous measures, proportions for other variables. *P* values based on the Welch *t*test for independent samples. HRT, hormone replacement therapy.

percentage density ($P = 0.0017$). Cases were found to have higher percentage density (mean ± standard deviation: 16.7 ± 14.3) than controls (14.6 ± 14.0). No significant difference was found for age, height, weight, body mass index, age at menarche, age at menopause or HRT usage.

Table S4 in Additional file 2 shows a list of 34 genes involved in the estrogen metabolic pathway and the corresponding number of SNPs examined for each gene. References are given for genes that have been examined in other studies for an association with mammographic density. Of the 239 SNPs analyzed, 11 SNPs were found to be significant at the 5% level (Table 2) - of which the smallest *P* value was 0.0019. Among six tagSNPs selected for the gene CYP11A1, five were found to be significant in the same direction. The associations in the single SNP analysis were moderate and would not survive correction for multiple SNP testing. In addition, the single-SNP *P* values showed no clear deviation from the null distribution, representing no association between SNPs and percentage density (Figure 2; see also Tables S1 to S3 in Additional file 1). None of the SNPs found to be nominally significant in our dataset were found to be significant in the CGEMS validation set (see Table S5 in Additional file 3). A second, independent validation carried out on the most significantly associated SNP (rs11638442) located within the CYP11A1 gene in 783 postmenopausal women with mammograms in the Mayo Clinic Breast Cancer Study yielded a *P* value of 0.88 (regression coefficient = -0.000507, 95% confidence interval = -0.07251 to 0.06237).

Since the estrogen metabolic SNPs examined have previously been associated with breast cancer risk, we estimated the correlation between regression coefficients of SNP effects on mammographic density and the odds ratios of SNP effects on breast cancer risk, in order to assess whether the SNPs act through mammographic density as an intermediate phenotype for breast cancer. No significant relationship was found between SNP effects on breast cancer risk and percentage density (Spearman's correlation rho = 0.0411, $P = 0.5268$). Pathway-based multi-SNP association analyses

Table 2 Significant SNPs in the estrogen metabolic pathway, corresponding regression coefficients and *P* values

SNP	Gene	Minor allele	MAF	*n*	Coefficient	SE	*P* value
rs11638442	CYP11A1	C	0.35	1,677	0.0557	0.0212	0.0088
rs16968478	CYP11A1	G	0.17	1,703	0.0575	0.0263	0.0293
rs2279357	CYP11A1	A	0.20	1,699	0.0511	0.0229	0.0260
rs2959003	CYP11A1	A	0.28	1,669	0.0582	0.0224	0.0094
rs2959008	CYP11A1	A	0.30	1,703	0.0475	0.0221	0.0315
rs2066485	HSD17B3	G	0.14	1,703	0.0668	0.0293	0.0230
rs7039978	HSD17B3	A	0.50	1,694	0.0632	0.0203	0.0019
rs1469908	NQO1	C	0.37	1,695	0.0472	0.0206	0.0223
rs17268974	STS	A	0.22	1,605	0.0503	0.0238	0.0349
rs2270112	STS	C	0.34	1,686	0.0485	0.0208	0.0197
rs707762	STS	A	0.40	1,687	0.0435	0.0205	0.0340

P values from a one degree of freedom likelihood ratio test. MAF, minor allele frequency; SE, standard error.

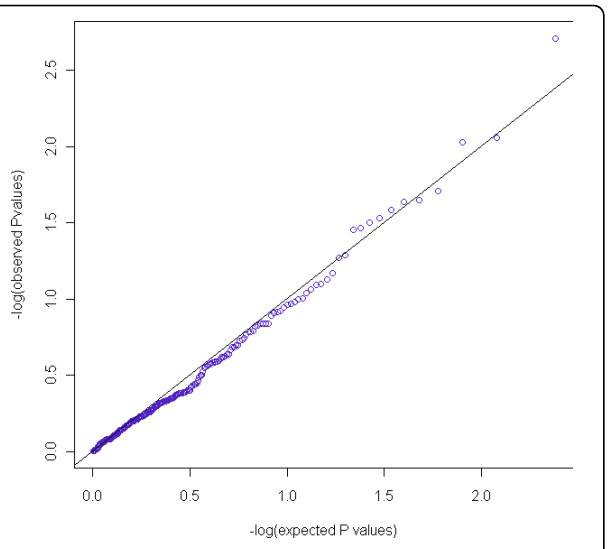


Figure 2 No association between SNPs and percentage density. $\log_{10}$ quantile quantile *P* value plots from single SNP trend tests of 239 SNPs in the estrogen metabolism pathway.

Table 3 Global genetic association tests between SNPs in the estrogen metabolic pathways and mammographic breast density

Pathway	Number of SNPs	*P* heterogeneity	*P* trend[a]
Whole pathway	239	0.840	0.507
Androgen synthesis	11	0.761	0.763
Androgen to estrogen conversion	120	0.587	0.715
Estrogen removal	134	0.834	0.872

[a]*P* values based on 5,000 permutations.

revealed no significant association between percentage density and genetic variations in the overall estrogen metabolic pathway, or any of the related subpathways (Table 3).

Discussion

Our study suggests there is no appreciable effect between genetic variants involved in estrogen metabolism and mammographic density. Neither the overall estrogen metabolic pathway nor the androgen synthesis, androgen-to-estrogen conversion and estrogen removal subpathways were found to be significantly associated with mammographic density. Single SNP markers with significant associations with mammographic density were not validated in two independent datasets.

In view of estrogen exposure being a major risk factor of postmenopausal breast cancer, and mammographic density being associated with several hormone-related factors such as body mass index (increased local estrogen conversion due to increased fatty tissue), HRT, and menopausal status, the estrogen metabolic pathway has been a candidate pathway for the search of genetic variants related to mammographic density. Most of the variants in the candidate breast cancer genes evaluated in previous studies, however, have been concluded to be only weak predictors of mammographic density [10]. Association findings have been both supported and contradicted [3]. As Boyd and colleagues have discussed [34], it is likely that hormone-related factors are responsible for only a small proportion of the wide variation in mammographic density. In addition, genetic variants involved in the estrogen metabolic pathway are generally investigated based on the premise that mammographic density is an intermediate and heritable risk factor of breast cancer [4]. There is, however, accumulating evidence that mammographic density may predispose to breast cancer risk through components largely independent of estrogen metabolism [35-37].

In our study, no correlation was found between the estimates of SNP effects on breast cancer risk and mammographic density, suggesting that the same SNPs associated with breast cancer risk are not directly correlated with mammographic density. Tamimi and colleagues reported that mammographic density and circulating sex steroid levels were independently associated with breast cancer risk in postmenopausal women [35]. In addition, Kerlikowske and colleagues found no correlation between mammographic density and bone mineral density [36], both of which have been suggested to be cumulative markers of elevated estrogen exposure. Dite and colleagues performed a similar study investigating the overlap between genetic determinants of mammographic density and bone mineral density, and reported a null finding [37]. Another finding in Kerlikowske and colleagues' study was that although mammographic density remained strongly associated with elevated breast cancer risk after adjustment for hormone-related factors, the effects of bone mineral density did not [36], suggesting that estrogen metabolism plays only a small role in the effects of mammographic density on breast cancer risk.

Many studies examining the effects of exogenous estrogen exposure are in agreement with the view that estrogen has limited effects on mammographic density. Very often, the combined estrogen plus progestin regimen was found to affect mammographic density more than the estrogen-only regimen [38-41], suggesting that progestins and not estrogens are responsible for increased mammographic density. Interestingly, mammographic density is also known to have no prognostic bearing on the estrogen receptor status of breast cancer tumors [42-44], thus corroborating an estrogen/estrogen receptor independent link. Another study conducted by Vachon and colleagues found no

influence of aromatase inhibitors (drugs that stop the production of estrogen in postmenopausal women) on mammographic density [45], further supporting this line of rationale.

Strengths of the present study include the large sample size and extensive coverage of SNPs in the estrogen metabolic pathway. In a review by Kelemen and colleagues, the authors summarized that previous genetic association studies exploring the relationship between the estrogen metabolic pathway and mammographic density had sample sizes ranging from between 232 and 1,260 women [3]. The number of loci involved in the estrogen metabolic pathway investigated in these studies was also limited to eight or less [3], while we examined 239 tagSNPs from 34 genes involved in the estrogen metabolic pathway. A second strength of the present study is the use of two independent populations for the validation of the associations found.

A limitation of the present work is that it includes different mammogram views across the different studies. The main study on Swedish women utilized the medio-lateral oblique view, while mammograms of the CGEMS and of the Mayo Clinic were taken using the CC view. Several studies, however, have shown correlation of densities from the medio-lateral oblique and CC views [46,47], and have shown that the different views yield similar associations with breast cancer [32]. In addition, the main focus of this study was on genetic determinants of mammographic density in postmenopausal women. Although no strong association was observed between SNPs in the estrogen metabolic pathway examined and mammographic density in postmenopausal women, whether the same lack of association between common genetic variation in the estrogen metabolism pathway and mammographic density is present in premenopausal women remains to be clarified.

Conclusions

As mammographic density is generally considered an intermediate phenotype for breast cancer, the identification of genes that influence mammographic density would play an important role in risk prediction of breast cancer prior to the start of mammography screenings and shed light on the mechanisms behind breast cancer carcinogenesis. Overall, there is no conclusive evidence that genetic variants in genes involved in the estrogen metabolic pathway are associated with mammographic density in postmenopausal women. This knowledge will be helpful for directing the focus of future studies to alternative pathways that may be responsible for a larger bulk of the genetic component of mammographic density.

Additional file 1: Tables S1 to S3. Table S1 presents a list of SNPs in the androgen synthesis subpathway and their corresponding regression coefficients and likelihood ratio test *P* values. Table S2 presents a list of SNPs in the androgen to estrogen conversion subpathway and their corresponding regression coefficients and likelihood ratio test *P* values. Table S3 presents a list of SNPs in the estrogen removal subpathway and their corresponding regression coefficients and likelihood ratio test *P* values.
Click here for file
[http://www.biomedcentral.com/content/supplementary/bcr2488 S1. DOC]

Additional file 2: Table S4. Table S4 presents genes containing polymorphisms within the estrogen metabolic pathway evaluated in relation to mammographic density.
Click here for file
[http://www.biomedcentral.com/content/supplementary/bcr2488 S2. DOC]

Additional file 3: Table S5. Table S5 presents validation results of significantly associated SNPs in the Nurses' Health Study (NHS) and the Mayo Clinic Breast Cancer Study (MBCS).
Click here for file
[http://www.biomedcentral.com/content/supplementary/bcr2488 S3. DOC]

Abbreviations

CC: craniocaudal; CGEMS: Cancer Genetic Markers of Susceptibility Project; HRT: hormone replacement therapy; SNP: single nucleotide polymorphism; tagSNP: tagging single nucleotide polymorphism.

Acknowledgements

The present work was supported by National Institutes of Health (RO1 CA58427), by Märit and Hans Rausing's Initiative Against Breast Cancer, and by the W81XWH 05 1 0314 Innovator Award, US Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs. JLi is a recipient of the A*STAR Graduate Scholarship. KH was supported by the Swedish Research Council (523 2006 972). KC was supported by the Swedish Cancer Society (5128 B07 01PAF). Validation work on CGEMS data by the Channing Laboratory was supported by the National Institutes of Health (P01 CA87969, R01 CA065725, R01 CA131332). Validation work at the Mayo Clinic was supported by R01 CA 128931 and R01 CA 122340.

Author details

[1]Karolinska Institutet, Department of Medical Epidemiology and Biostatistics, Box 281, 171 77 Stockholm, Sweden. [2]Human Genetics, Genome Institute of Singapore, 60 Biopolis Street, Singapore 138672, Singapore. [3]Channing Laboratory, Brigham and Women's Hospital, 181 Longwood Avenue, Boston, MA 02115, USA. [4]Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. [5]Program in Genetic and Molecular Epidemiology, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. [6]Department of Health Sciences Research, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905, USA. [7]Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905, USA. [8]Department of Hygiene, Epidemiology and Medical Statistics, University of Athens Medical School, 75 Mikras Asias Str, Goudi, Athens 115 27, Greece.

Authors' contributions

JLi participated in the study design, carried out the analyses and drafted the manuscript. LE digitized and obtained readings for the mammograms. RMT, SL, DJH, CMV, FJC and CGS contributed to the validation of this study. PL coordinated the Innovator project which contributed data on birthweight and mammographic density. JLi, KH, KC, JLiu and PH conceived of the study, and participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Received: 29 December 2009 Revised: 1 March 2010
Accepted: 9 March 2010 Published: 9 March 2010

References

1. Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ: **Mammographic densities and breast cancer risk.** *Cancer Epidemiol Biomarkers Prev* 1998, **7**:1133 1144.
2. Vachon CM, van Gils CH, Sellers TA, Ghosh K, Pruthi S, Brandt KR, Pankratz VS: **Mammographic density, breast cancer risk and risk prediction.** *Breast Cancer Res* 2007, **9**:217.
3. Kelemen LE, Sellers TA, Vachon CM: **Can genes for mammographic density inform cancer aetiology?.** *Nat Rev Cancer* 2008, **8**:812 823.
4. Boyd NF, Dite GS, Stone J, Gunasekara A, English DR, McCredie MR, Giles GG, Tritchler D, Chiarelli A, Yaffe MJ, Hopper JL: **Heritability of mammographic density, a risk factor for breast cancer.** *N Engl J Med* 2002, **347**:886 894.
5. Stone J, Gurrin LC, Byrnes GB, Schroen CJ, Treloar SA, Padilla EJ, Dite GS, Southey MC, Hayes VM, Hopper JL: **Mammographic density and candidate gene variants: a twins and sisters study.** *Cancer Epidemiol Biomarkers Prev* 2007, **16**:1479 1484.
6. Boyd NF, Martin LJ, Rommens JM, Paterson AD, Minkin S, Yaffe MJ, Stone J, Hopper JL: **Mammographic density: a heritable risk factor for breast cancer.** *Methods Mol Biol* 2009, **472**:343 360.
7. Boyd NF, Martin LJ, Yaffe MJ, Minkin S: **Mammographic density: a hormonally responsive risk factor for breast cancer.** *J Br Menopause Soc* 2006, **12**:186 193.
8. Duffy SW, Jakes RW, Ng FC, Gao F: **Interaction of dense breast patterns with other breast cancer risk factors in a case control study.** *Br J Cancer* 2004, **91**:233 236.
9. Martin LJ, Boyd NF: **Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence.** *Breast Cancer Res* 2008, **10**:201.
10. Haiman CA, Hankinson SE, De Vivo I, Guillemette C, Ishibe N, Hunter DJ, Byrne C: **Polymorphisms in steroid hormone pathway genes and mammographic density.** *Breast Cancer Res Treat* 2003, **77**:27 36.
11. Maskarinec G, Lurie G, Williams AE, Le Marchand L: **An investigation of mammographic density and gene variants in healthy women.** *Int J Cancer* 2004, **112**:683 688.
12. Lord SJ, Mack WJ, Berg Van Den D, Pike MC, Ingles SA, Haiman CA, Wang W, Parisky YR, Hodis HN, Ursin G: **Polymorphisms in genes involved in estrogen and progesterone metabolism and mammographic density changes in women randomized to postmenopausal hormone therapy: results from a pilot study.** *Breast Cancer Res* 2005, **7**:R336 R344.
13. Chambo D, Kemp C, Costa AM, Souza NC, Guerreiro da Silva ID: **Polymorphism in CYP17, GSTM1 and the progesterone receptor genes and its relationship with mammographic density.** *Braz J Med Biol Res* 2009, **42**:323 329.
14. Haiman CA, Bernstein L, Berg D, Ingles SA, Salane M, Ursin G: **Genetic determinants of mammographic density.** *Breast Cancer Res* 2002, **4**:R5.
15. van Duijnhoven FJ, Bezemer ID, Peeters PH, Roest M, Uitterlinden AG, Grobbee DE, van Gils CH: **Polymorphisms in the estrogen receptor alpha gene and mammographic density.** *Cancer Epidemiol Biomarkers Prev* 2005, **14**:2655 2660.
16. Wedren S, Lovmar L, Humphreys K, Magnusson C, Melhus H, Syvanen AC, Kindmark A, Landegren U, Fermer ML, Stiger F, Persson I, Baron J, Weiderpass E: **Oestrogen receptor alpha gene haplotype and postmenopausal breast cancer risk: a case control study.** *Breast Cancer Res* 2004, **6**:R437 R449.
17. Tamimi RM, Eriksson L, Lagiou P, Czene K, Ekbom A, Hsieh CC, Adami HO, Trichopoulos D, Hall P: **Birth weight and mammographic density among postmenopausal women in Sweden.** *Int J Cancer* **126**:985 991.
18. Byng JW, Boyd NF, Little L, Lockwood G, Fishell E, Jong RA, Yaffe MJ: **Symmetry of projection in the quantitative analysis of mammographic images.** *Eur J Cancer Prev* 1996, **5**:319 327.
19. Qin ZS, Niu T, Liu JS: **Partition ligation expectation maximization algorithm for haplotype inference with single nucleotide polymorphisms.** *Am J Hum Genet* 2002, **71**:1242 1247.
20. Stram DO, Haiman CA, Hirschhorn JN, Altshuler D, Kolonel LN, Henderson BE, Pike MC: **Choosing haplotype tagging SNPS based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the Multiethnic Cohort Study.** *Hum Hered* 2003, **55**:27 36.
21. Tyrer J, Pharoah PD, Easton DF: **The admixture maximum likelihood test: a novel experiment wise test of association between disease and multiple SNPs.** *Genet Epidemiol* 2006, **30**:636 643.
22. R Development Core Team: *R: A Language and Environment for Statistical Computing* Austria: R Foundation for Statistical Computing 2008.
23. Everett BM, Kurth T, Buring JE, Ridker PM: **The relative strength of C reactive protein and lipid levels as determinants of ischemic stroke compared with coronary heart disease in women.** *J Am Coll Cardiol* 2006, **48**:2235 2242.
24. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, Chatterjee N, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Hayes RB, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover RN, Thomas G, Chanock SJ: **A genome wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.** *Nat Genet* 2007, **39**:870 874.
25. Colditz GA, Hankinson SE: **The Nurses' Health Study: lifestyle and health among women.** *Nat Rev Cancer* 2005, **5**:388 396.
26. Hankinson SE, Willett WC, Manson JE, Colditz GA, Hunter DJ, Spiegelman D, Barbieri RL, Speizer FE: **Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women.** *J Natl Cancer Inst* 1998, **90**:1292 1299.
27. Tamimi RM, Cox DG, Kraft P, Pollak MN, Haiman CA, Cheng I, Freedman ML, Hankinson SE, Hunter DJ, Colditz GA: **Common genetic variation in IGF1, IGFBP 1, and IGFBP 3 in relation to mammographic density: a cross sectional study.** *Breast Cancer Res* 2007, **9**:R18.
28. Li Y, Willer C, Sanna S, Abecasis G: **Genotype imputation.** *Annu Rev Genomics Hum Genet* 2009, **10**:387 406.
29. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM: **GenABEL: an R library for genome wide association analysis.** *Bioinformatics* 2007, **23**:1294 1296.
30. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odefrey F, Shen CY, Wu PE, Wang HC, Eccles D, Evans DG, Peto J, Fletcher O, *et al*: **Genome wide association study identifies novel breast cancer susceptibility loci.** *Nature* 2007, **447**:1087 1093.
31. Cox A, Dunning AM, Garcia Closas M, Balasubramanian S, Reed MW, Pooley KA, Scollen S, Baynes C, Ponder BA, Chanock S, Lissowska J, Brinton L, Peplonska B, Southey MC, Hopper JL, McCredie MR, Giles GG, Fletcher O, Johnson N, dos Santos Silva I, Gibson L, Bojesen SE, Nordestgaard BG, Axelsson CK, Torres D, Hamann U, Justenhoven C, Brauch H, Chang Claude J, Kropp S, *et al*: **A common coding variant in CASP8 is associated with breast cancer risk.** *Nat Genet* 2007, **39**:352 358.
32. Vachon CM, Brandt KR, Ghosh K, Scott CG, Maloney SD, Carston MJ, Pankratz VS, Sellers TA: **Mammographic breast density as a general marker of breast cancer risk.** *Cancer Epidemiol Biomarkers Prev* 2007, **16**:43 49.
33. Boyd NF, Stone J, Martin LJ, Jong R, Fishell E, Yaffe M, Hammond G, Minkin S: **The association of breast mitogens with mammographic densities.** *Br J Cancer* 2002, **87**:876 882.
34. Boyd NF, Rommens JM, Vogt K, Lee V, Hopper JL, Yaffe MJ, Paterson AD: **Mammographic breast density as an intermediate phenotype for breast cancer.** *Lancet Oncol* 2005, **6**:798 808.
35. Tamimi RM, Byrne C, Colditz GA, Hankinson SE: **Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women.** *J Natl Cancer Inst* 2007, **99**:1178 1187.
36. Kerlikowske K, Shepherd J, Creasman J, Tice JA, Ziv E, Cummings SR: **Are breast density and bone mineral density independent risk factors for breast cancer?** *J Natl Cancer Inst* 2005, **97**:368 374.
37. Dite GS, Wark JD, Giles GG, English DR, McCredie MR, Hopper JL: **Is there overlap between the genetic determinants of mammographic density and bone mineral density?** *Cancer Epidemiol Biomarkers Prev* 2005, **14**:2266 2268.
38. Kaewrudee S, Anuwutnavin S, Kanpittaya J, Soontrapa S, Sakondhavat C: **Effect of estrogen progestin and estrogen on mammographic density.** *J Reprod Med* 2007, **52**:513 520.

39. McTiernan A, Martin CF, Peck JD, Aragaki AK, Chlebowski RT, Pisano ED, Wang CY, Brunner RL, Johnson KC, Manson JE, Lewis CE, Kotchen JM, Hulka BS: **Estrogen plus progestin use and mammographic density in postmenopausal women: Women's Health Initiative randomized trial.** *J Natl Cancer Inst* 2005, **97**:1366 1376.
40. Topal NB, Ayhan S, Topal U, Bilgin T: **Effects of hormone replacement therapy regimens on mammographic breast density: the role of progestins.** *J Obstet Gynaecol Res* 2006, **32**:305 308.
41. McTiernan A, Chlebowski RT, Martin C, Peck JD, Aragaki A, Pisano ED, Wang CY, Johnson KC, Manson JE, Wallace RB, Vitolins MZ, Heiss G: **Conjugated equine estrogen influence on mammographic density in postmenopausal women in a substudy of the women's health initiative randomized trial.** *J Clin Oncol* 2009, **27**:6135 6143.
42. Ghosh K, Brandt KR, Sellers TA, Reynolds C, Scott CG, Maloney SD, Carston MJ, Pankratz VS, Vachon CM: **Association of mammographic density with the pathology of subsequent breast cancer among postmenopausal women.** *Cancer Epidemiol Biomarkers Prev* 2008, **17**:872 879.
43. Chen JH, Hsu FT, Shih HN, Hsu CC, Chang D, Nie K, Nalcioglu O, Su MY: **Does breast density show difference in patients with estrogen receptor positive and estrogen receptor negative breast cancer measured on MRI?** *Ann Oncol* 2009, **20**:1447 1449.
44. Ziv E, Tice J, Smith Bindman R, Shepherd J, Cummings S, Kerlikowske K: **Mammographic density and estrogen receptor status of breast cancer.** *Cancer Epidemiol Biomarkers Prev* 2004, **13**:2090 2095.
45. Vachon CM, Ingle JN, Suman VJ, Scott CG, Gottardt H, Olson JE, Goss PE: **Pilot study of the impact of letrozole vs. placebo on breast density in women completing 5 years of tamoxifen.** *Breast* 2007, **16**:204 210.
46. Ursin G, Hovanessian Larsen L, Parisky YR, Pike MC, Wu AH: **Greatly increased occurrence of breast cancers in areas of mammographically dense tissue.** *Breast Cancer Res* 2005, **7**:R605 R608.
47. Kim SJ, Moon WK, Cho N, Cha JH, Kim SM, Im JG: **Computer aided detection in digital mammography: comparison of craniocaudal, mediolateral oblique, and mediolateral views.** *Radiology* 2006, **241**:695 701.

doi:10.1186/bcr2488
Cite this article as: Li *et al.*: Genetic variation in the estrogen metabolic pathway and mammographic density as an intermediate phenotype of breast cancer. *Breast Cancer Research* 2010 **12**:R19.

Eur J Epidemiol (2011) 26:39 44
DOI 10.1007/s10654 010 9528 3

DEVELOPMENTAL EPIDEMIOLOGY

Energy intake during pregnancy in relation to offspring gender by maternal height

Pagona Lagiou · Evangelia Samoli · Loren Lipworth · Areti Lagiou · Fang Fang · Marta Rossi · Biao Xu · Guo-Pei Yu · Hans-Olov Adami · Chung-Cheng Hsieh · Dimitrios Trichopoulos

Received: 24 August 2010 / Accepted: 18 November 2010 / Published online: 30 November 2010

Abstract Male newborns are somewhat heavier than female ones and it has been reported, in a Caucasian population, that mothers carrying boys have higher energy intake during pregnancy compared to those carrying girls. In the context of a prospective study comprising 150 Caucasian women in Boston, USA and 243 Asian women in Shanghai China, energy intake at the second trimester of pregnancy was estimated based on center-specific food frequency questionnaires. There was a significant interaction ($P = 0.01$) of maternal height with offspring gender with respect to maternal daily energy intake. Among taller women, male gender of the offspring was associated with higher maternal energy intake (difference 341 kcal/day, 95% Confidence Interval 77 604; $P = 0.01$), whereas among shorter women, no significant association existed between offspring gender and maternal daily energy intake (difference −213 kcal/day, 95% Confidence Interval −479 to 54; $P = 0.12$). Our findings indicate that the higher somatic growth potential of boys in intrauterine life is realized only when there are no constrains imposed by maternal anthropometry and it is, then, associated with higher maternal energy intake during pregnancy.

Keywords Birth weight · Energy intake · Maternal height · Offspring gender · Pregnancy

Introduction

It is established and appears consistent across populations that male newborns are somewhat heavier than female ones [1]. In order to examine whether the difference is

P. Lagiou (✉) · H. O. Adami · C. C. Hsieh · D. Trichopoulos
Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA
e mail: plagiou@hsph.harvard.edu

P. Lagiou · E. Samoli
Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, University of Athens, 75 M. Asias Street, Goudi, 115 27 Athens, Greece

L. Lipworth
International Epidemiology Institute, 1455 Research Blvd, Suite 550, Rockville, MD 20850, USA

A. Lagiou
Faculty of Health Professions, Athens Technological Educational Institute, 274 Thivon Avenue, 122 41 Athens, Greece

F. Fang · H. O. Adami · D. Trichopoulos
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, 171 77 Stockholm, Sweden

M. Rossi
Laboratorio di Epidemiologia, Istituto di Ricerche Farmacologiche ‘‘Mario Negri’’, Via Eritrea, 62 20157 Milan, Italy

B. Xu
Department of Epidemiology, School of Public Health, Fudan University, 200023 Shanghai, China

G. P. Yu
Biostatistics and Epidemiology Service, The New York Eye and Ear Infirmary, 310 14th Street, New York, NY 10003, USA

C. C. Hsieh
Division of Biostatistics and Epidemiology, UMass Cancer Center, University of Massachusetts Medical School, 364 Plantation Street, LRB 427, Worcester, MA 01605, USA

P. Lagiou · D. Trichopoulos
Bureau of Epidemiologic Research, Academy of Athens, Athens, Greece

accounted for by higher energy intake, or more efficient energy utilization by a pregnant woman carrying a male rather than a female embryo, a study among Caucasian women in Boston, USA was undertaken. In that study, the only one on this topic in the literature, we found that pregnant women carrying a boy have an approximately 10% higher energy intake compared to those carrying a girl [2]. That study relied on an international prospective study on pregnancy hormones and outcomes among women in Boston, USA and Shanghai, China [3]. Because in earlier studies relying on this cohort the database for the Chinese diet was not available, only data on dietary intakes for the US women were reported [3, 4]. With procession of the dietary analyses on the Shanghai China data, we report now results based on data from both cohorts. We were particularly interested to examine the role of maternal height, which is different between Chinese and Caucasian women and is a crucial determinant of birth weight [5, 6], in the association of gender offspring with maternal energy intake.

Materials and methods

Subjects

Between March 1994 and October 1995, authorized health professionals met all pregnant women coming for their first routine prenatal visit to the collaborating maternity clinics of the Beth Israel hospital in Boston, USA and the Shanghai Medical University in China [3]. They ascertained the woman's eligibility to participate, explained to her the objectives of the study and obtained informed consent. A total of 402 Caucasian women in Boston, USA and 424 Asian women in Shanghai, China were identified. Eligibility criteria included age less than 40 years (to minimize selection bias generated by the unusual occurrence of pregnancies beyond that age and the higher proportion of assisted reproduction in that age group), a maximum parity of two, absence of a prior diagnosis of diabetes mellitus or thyroid disease, no hormonal medication during the index pregnancy and no known fetal abnormality. The maximum parity of two criterion was imposed by the one-child policy implemented in China.

Of the 402 eligible women in Boston, USA, 77 refused to participate in one or more aspects of the study, 9 were subsequently excluded because of a spontaneous or induced abortion in the index pregnancy, 2 were excluded because of twin birth, whereas 10 were lost to follow-up after the initial meeting. Of the remaining 304 women, we excluded 35 women with gestational age below 37 or above 42 weeks, 16 additional women with pregnancy toxemia and another 38 women with missing information on essential (e.g., age) or multiple covariates. We have considered imputations for these women but we have eventually opted for an assumption-free approach rather for a marginal increase in the statistical power of the study. For the present analyses we excluded 65 women with daily energy intake below 1,600 or above 6,000 kcal, which correspond to the lower 5% and upper 90% of energy intake among Chinese women. Eventually, 150 Caucasian women were considered in this study.

Of the 424 eligible women in Shanghai, China, 15 refused to participate in one or more aspects of the study, 2 women were excluded owing to induced abortion in the index pregnancy and another 2 because of twin birth, 7 women were lost to follow-up and 44 women were excluded because they had gestation duration outside the range of 37 42 weeks inclusive. There were no Asian women with preeclampsia. Another 59 women were excluded because of missing information on essential (e.g., age) or multiple covariates. For the present analyses we excluded 52 women with daily energy intake below 1,600 or above 6,000 kcal (which, as indicated, correspond to the lower 5% and upper 90% of energy intakes among Chinese women). Eventually, 243 Asian women were considered in this study.

The study was approved by the Institutional Review Boards of the Beth Israel Hospital, Shanghai Medical University and Harvard School of Public Health.

Baseline data and measurements

Baseline socio-demographic and lifestyle information was recorded in interviews at the 16th and the 27th gestational week visit of the women to the clinic. Information about medical conditions, maternal anthropometry and weight gain by the 16th and the 27th gestational week was abstracted from the women's medical records. At delivery, additional information concerning the newborn, including duration of gestation and birth size parameters, was recorded, as previously reported [3].

Dietary intakes during the second trimester were assessed through an extensive semi-quantitative food frequency questionnaire (FFQ) filled in by the women and checked for accuracy and completeness by trained interviewers during the 2nd visit at the maternity clinic around the 27th gestational week. For baseline socio-demographic and lifestyle variables, the information on both visits were taken into account. In the dietary questionnaires there were very few instances (<1%) where consumption information was not reported and we have assumed that the corresponding food item was not consumed. For women in Boston, USA, the FFQ was identical to the one used and validated in the Nurses' Health Study [7]. Intake of energy and energy-generating nutrients were calculated from the dietary data using the standard software used in the Nurses'

Health Study [8]. For women in Shanghai, China, the FFQ was reformulated so as to be compatible with the local dietary patterns and covered approximately 135 food and beverage items, but it was not formally validated. Intake of energy was calculated from the dietary data using the food composition tables developed by the Chinese Institute of Nutrition and Food Safety [9, 10].

Information concerning the study protocol and implementation has been published [3].

Statistical analyses

Statistical analyses were conducted using the SPSS statistical package (Statistical Package for Social Sciences v. 16.0, Chicago, Illinois, USA). We distributed

Caucasian and Chinese women by maternal offspring characteristics and, for quantitative variables, we estimated the mean values and standard deviations. Missing pre-pregnancy maternal weight values (2 in each center) were replaced by the respective center-specific age and height-adjusted mean values, whereas missing maternal weight gain values (12 in Boston, USA and 5 in Shanghai, China) were replaced by the respective center-specific mean values for body mass index (BMI) values above or below 22, as appropriate.

Initially, in each center, we applied multiple regression models to estimate the effect of the gender of the fetus (male vs female) on maternal daily energy intake, controlling for maternal age (categorically; less than 30 years of age, 30 34 and 35+ years), educational level (categorically; high school graduate, college graduate, higher, unspecified), parity (only in the models applied for Boston, USA, women since almost all Chinese women had parity 1), duration of gestation (continuously; in weeks), maternal height (continuously, in centimeters), pre-pregnancy weight (continuously, in kg) and weight gain during pregnancy (continuously, in kg). We then ran regression models per center separately, as well as overall controlling for center, adding to the covariates indicated above an interaction term between gender of offspring and maternal height. Finally, we applied multiple regression models per center separately, as well as overall controlling for center, for women of short and tall stature according to the center-specific median cut off, i.e., 165 cm for Caucasian women and 160 cm for Chinese women.

Results

Table 1 shows characteristics of women in the two centers. Women in Shanghai, China in comparison to women in Boston, USA, were younger, mostly primiparae, of lower education, with lower stature and pre-pregnancy body weight and gained less weight during pregnancy ($P < 0.001$ in all these comparisons). Daily energy intake during the second trimester of pregnancy, estimated on the basis of reported food consumption, was substantially higher among women in Shanghai, China ($P < 0.001$); it is noted, however, that the FFQ used in Shanghai, China collected more detailed information and, hence, was more likely to lead to overestimation [11, 12], than the questionnaire used in Boston, USA. The sex ratio was also significantly different between the two centers ($P = 0.02$). All indicated variables were adjusted for in the analyses.

We regressed separately for women in Boston, USA and Shanghai, China daily energy intake on maternal age, parity (only in Boston, USA), education, pre-pregnancy body weight, maternal height, weight gain until the 27th gestational week, exact duration of gestation and offspring gender. In Boston, USA, only gender of the offspring was a marginally significant predictor of maternal energy intake, male vs female gender being associated with a daily energy intake higher by 173 kcal (95% confidence interval (CI): −1 to 347, $P = 0.05$). In Shanghai, China, gender of the offspring was unrelated to energy intake (regression coefficient of boys vs girls -54 kcal/day (95%CI: −351 to 244), $P = 0.72$), and only maternal height was a marginally significant predictor of maternal energy intake, with an increase of 364 kcal/day per 10 cm increment in maternal height (95%CI: 4 724, $P = 0.05$). We repeated the model for both centers combined, controlling for center. Maternal height remained a significant predictor ($P = 0.04$) of maternal energy intake, with an increase of 209 kcal/day per 10 cm increment in maternal height (95%CI: 13 406), whereas the association with offspring gender was non-significant (regression coefficient of boys vs girls 41 kcal/day (95%CI: −147 to 229), $P = 0.67$).

Because gender was a significant predictor of maternal energy intake only in Boston, USA, whereas maternal height was a significant predictor of maternal energy intake only in Shanghai, China, and because both maternal height and offspring gender were significantly different between the two centers, we ran models regressing maternal daily energy intake on the predictor variables previously indicated, but also including an interaction term for offspring gender by maternal height (continuous), separately in the two centers, as well as in both centers combined, controlling for center. The results are presented in Table 2. For maternal height and gender of offspring with respect to daily energy intake, there was a significant interaction over both centers and in Boston, USA and a suggestive one in the same direction in Shanghai, China. These results indicate that, among taller women, male gender of the offspring is associated with higher maternal energy intake, whereas among shorter women, no

Table 1 Distribution of pregnant women and offspring gender in Boston, USA and Shanghai, China by demographic and somatometric characteristics

	Boston, USA (*n* 150)		Shanghai, China (*n* 243)	
Maternal age (in years)	*n*	%	*N*	%
<30	38	25.3	213	87.7
30 34	97	64.7	22	9.1
35 39	15	10.0	8	3.3
Parity				
1	85	56.7	239	98.4
2+	65	43.3	4	1.6
Education				
High school graduate	11	7.3	110	45.3
College graduate	69	46.0	82	33.7
Higher	70	46.7	30	12.4
Unspecified	0	0	21	8.6
Maternal height (in meters)				
1.54	14	9.4	29	11.9
1.55 1.59	16	10.7	74	30.5
1.60 1.64	34	22.7	97	39.9
1.65 1.69	42	28.0	37	15.2
1.70 1.74	29	19.3	6	2.5
1.75+	15	10.0	0	0.0
Maternal pre pregnancy weight (in kg)				
49	25	16.7	97	39.9
50 59	68	45.3	125	51.5
60 69	39	26.0	19	7.8
70+	18	12.0	2	0.8
Maternal weight gain until the 27th gestational week (in kg)	11.4	4.1	8.6	4.3
Duration of gestation (in weeks)	40.1	1.2	40.0	1.1
Daily energy intake (in kcal)	2,328	541	3,321	1,105
Gender of offspring				
Male	73	48.7	147	60.5
Female	77	51.3	96	39.5

For continuous variables mean (SD) for categorical *n* (%)

Women with gestational age between 37 and 42 weeks, no eclampsia and daily energy intake between 1,600 and 6,000 kcal

significant association exists between offspring gender and maternal daily energy intake. We repeated the analysis in the Boston, USA, cohort taking into account only women with parity 1 and the critical statistic, the regression coefficient among women with high stature in Boston, USA, had, if anything, slightly increased from 348 to 365 kcal/day.

Since women in Shanghai, China, were generally shorter than women in Boston, USA (Table 1), one would expect, on the basis of the results in Table 2, that the association of maternal daily energy intake with offspring gender would be more pronounced among Caucasian women in Boston, USA than among Chinese women in Shanghai, China, as indeed they were. In line with these results, birth weight and birth length of offspring were substantially different by offspring gender in Boston, USA but not so in Shanghai, China (Table 3).

Discussion

In a prospective study of pregnant women in Boston, USA and Shanghai, China, we found that among women of shorter stature, offspring gender was not significantly associated with maternal daily energy intake. In contrast, among taller women, carrying a male, rather than a female, offspring was significantly associated with higher maternal daily energy intake.

To our knowledge, there has been only one previous study examining gender of offspring in relation to maternal energy intake, and this relied on the Boston, USA subsample of the present investigation [2]. In that study, among the generally tall Caucasian women in Boston, USA, a significantly higher maternal energy intake was found among women carrying a boy rather than a girl, but no attempt was made to study interaction of offspring

Table 2 Regression derived mean difference (b) in daily energy intake of pregnant women carrying a boy rather than a girl (baseline), by maternal height and study center

Maternal stature	b (kcal)	95% CI	*P* value	*P* for interaction[a]
Boston, USA (boys vs girls)				
≤165 cm (*n* 64)	6.2	235 to 222	0.96	0.03
>165 cm (*n* 86)	348	82 to 614	0.01	
Shanghai, China (boys vs girls)				
≤160 cm (*n* 148)	332	716 to 52	0.09	0.16
>160 cm (*n* 95)	275	216 to 766	0.27	
Both centers (boys vs girls)				
Shorter stature (*n* 212)[b]	213	479 to 54	0.12	0.01
Taller stature (*n* 181)[b]	341	77 to 604	0.01	

Controlling for maternal age (categorically), educational level (categorically), parity (for Boston, USA only), duration of gestation (continu ously), pre pregnancy weight (continuously), weight gain during pregnancy (continuously) and maternal height (continuously)

[a] *P* for interaction of offspring gender by maternal height (continuous) with respect to maternal daily energy intake

[b] Definition of taller versus shorter is center specific as indicated. Controlling for center, maternal age (categorically), educational level (categorically), parity, duration of gestation (continuously), pre pregnancy weight (continuously), weight gain during pregnancy (continuously) and maternal height (continuously)

Table 3 Mean values of somatometric characteristics of the off spring and energy intake among pregnant women by gender of the offspring in Boston, USA and Shanghai, China

	Boys	Girls	*P* value[a]
Boston, USA			
Birth weight (gr)	3,653	3,520	0.03
Birth length (cm)	51.2	50.0	0.002
Energy intake (kcal/day)	2,422	2,239	0.05
Shanghai, China			
Birth weight (gr)	3,436	3,414	0.71
Birth length (cm)	49.8	50.2	0.15
Energy intake (kcal/day)	3,311	3,336	0.76

[a] Comparing boys vs girls through the non parametric Mann Whit ney test

gender and maternal height with regard to maternal energy intake.

Energy requirements from infancy to adulthood are higher for males compared to females [13]. Our findings indicate that this pattern prevails in intrauterine life also, but in order to be manifested there should be minimal or no constrains imposed by maternal anthropometry on the fetal growth potential. Thus, among mothers of taller stature, birth size is generally bigger [3, 5, 14, 15], and differences in energy intakes during pregnancy between mothers carrying a boy compared to a girl, as well as differences in birth size between male and female offspring are pronounced. In contrast, among mothers of shorter stature, not only is birth size smaller, but differences in energy intake during pregnancy between mothers carrying a boy compared to a girl, as well as differences in birth size between male and female offspring are also minimized.

Strengths of our study are its prospective design, reliance of a uniform protocol in both centers and a reasonably large sample size. A weakness of the study was the use of different FFQs in the two centers which was imposed by the need to accommodate the characteristics of local diets. Indeed, the Chinese FFQ recorded detailed information, which could lead to overestimation of total energy intake [11, 12]. Although we excluded unreasonable high values of energy intake and even though many Chinese women were from rural areas, the level of energy intake in the Shanghai, China cohort was still unusually high. However, overestimation of energy intake is unlikely to be associated with measurement of maternal height or gender of the offspring and the systematic difference in reported energy intake between Boston, USA, and Shanghai, China, does not affect the reported results, which were either studied within center, or after controlling for center. Several exclusions were necessary, but these were generally undertaken for technical reasons, which are unlikely to have introduced bias. In some instances exclusions were made because of missing values (38 women in Boston, USA, that is 9.5% of the total, and 52 in Shanghai, China that is 12.3% of the total), but we have decided not to undertake imputation, opting for more conservative findings rather for marginal increase of statistical power. The Chinese FFQ was modeled according to the American FFQ, but it was not independently validated. This would be compatible with increased misclassification and effect attenuation, but it could not bias the results away from the null, since any errors in exposure ascertainment are not associated with possible errors in the later recorded outcome variables.

In conclusion, we have found evidence that among women of taller stature, but not among women of shorter stature, offspring gender is significantly associated with maternal daily energy intake. These findings need to be replicated, but we interpret the existing evidence as indicating that the higher somatic growth potential of boys in intrauterine life may be realized only when there are no constrains imposed by maternal anthropometry and is, then, associated with higher maternal energy intake during pregnancy.

Acknowledgments This work was supported by the W81XWH 05 1 0314 Innovator Award, US. Department of Defense Breast Cancer Research Program, Office of the Congressionally Directed Medical Research Programs.

References

1. Van Vliet G, Liu S, Kramer MS. Decreasing sex difference in birth weight. Epidemiology. 2009;20:622.
2. Tamimi RM, Lagiou P, Mucci LA, et al. Average energy intake among pregnant women carrying a boy compared with a girl. BMJ. 2003;326:1245 6.
3. Lipworth L, Hsieh C c, Wide L, et al. Maternal pregnancy hor mone levels in an area with a high incidence (Boston, USA) and in an area with a low incidence (Shanghai, China) of breast cancer. Brit J Cancer. 1999;79:7 12.
4. Lagiou P, Lagiou A, Samoli E, Hsieh CC, Adami HO, Tricho poulos D. Diet during pregnancy and levels of maternal preg nancy hormones in relation to the risk of breast cancer in the offspring. Eur J Cancer Prev. 2006;15(1):20 6.
5. Lagiou P, Samoli E, Lagiou A, et al. Maternal height, pregnancy estriol and birth weight in reference to breast cancer risk in Boston and Shanghai. Int J Cancer. 2005;117:494 8.
6. Xue F, Willett WC, Rosner BA, et al. Parental characteristics as predictors of birthweight. Hum Reprod. 2008;23:168 77.
7. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol. 1985;122:51 65.
8. Romieu I, Stampfer MJ, Stryker WS, et al. Food predictors of plasma beta carotene and alpha tocopherol: validation of a food frequency questionnaire. Am J Epidemiol. 1990;131(5):864 76.
9. Yang Y, Wang G, Pan X, editors. China food composition table 2002 (Book 1). Institute of Nutrition and Food Safety, China CDC. China: Peking University Medical Press; 2002. p. 1 395.
10. Yang Y, Wang G, Pan X, editors. China food composition table 2004 (Book 2). Institute of Nutrition and Food Safety, China CDC. China: Peking University Medical Press; 2004. p. 1 351.
11. Serdula M, Byers T, Coates R, et al. Assessing consumption of high fat foods: the effect of grouping foods into single questions. Epidemiology. 1992;3:503 8.
12. Nelson M, Bingham S. Assessment of food consumption and nutrient intake. In: Margetts BM, Nelson M, editors. Design concepts in nutritional epidemiology. 2nd ed. New York: Oxford University Press; 2000. p. 147.
13. FAO Food and Nutrition Technical Report Series. Human energy requirements. Report of a Joint FAO/WHO/UNU Expert Con sultation. Rome, 17 24 October 2001. p. ix + 96.
14. Cawley RH, McKeown T, Record RG. Parental stature and birthweight. Am J Hum Genet. 1954;6:448 56.
15. Emanuel I, Kimpo C, Moceri V. The association of maternal growth and socio economic measures with infant birthweight in four ethnic groups. Int J Epidemiol. 2004;33:1236 42.